

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 25.2632 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-2

Perfect score: 31

Sequence: 1 NYGVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 31 | 100.0 | 5 | 2 AAW08938 | AAW08938 CDR1 from |
| 2 | 31 | 100.0 | 5 | 3 AAY59311 | AAY59311 Heavy cha |
| 3 | 31 | 100.0 | 5 | 4 AAB37952 | AAB37952 Anti-EGFR |
| 4 | 31 | 100.0 | 5 | 5 AAU77784 | AAU77784 Mouse hea |
| 5 | 31 | 100.0 | 19 | 6 ABO32615 | ABO32615 Secreted |
| 6 | 31 | 100.0 | 107 | 4 AAM93336 | AAM93336 Human pol |
| 7 | 31 | 100.0 | 116 | 2 AAW05823 | AAW05823 Humanised |
| 8 | 31 | 100.0 | 117 | 3 AAG33864 | AAG33864 Arabidops |
| 9 | 31 | 100.0 | 119 | 2 AAW08953 | AAW08953 Heavy cha |
| 10 | 31 | 100.0 | 119 | 2 AAW08950 | AAW08950 Heavy cha |
| 11 | 31 | 100.0 | 119 | 2 AAW08952 | AAW08952 Heavy cha |
| 12 | 31 | 100.0 | 119 | 2 AAW08951 | AAW08951 Heavy cha |
| 13 | 31 | 100.0 | 119 | 2 AAW08954 | AAW08954 Heavy cha |
| 14 | 31 | 100.0 | 119 | 3 AAY90818 | AAY90818 260F9 hyb |
| 15 | 31 | 100.0 | 130 | 3 AAB00210 | AAB00210 Human seq |
| 16 | 31 | 100.0 | 138 | 2 AAW08947 | AAW08947 Heavy cha |
| 17 | 31 | 100.0 | 138 | 2 AAW08944 | AAW08944 Heavy cha |
| 18 | 31 | 100.0 | 138 | 2 AAW08942 | AAW08942 Heavy cha |
| 19 | 31 | 100.0 | 141 | 3 AAY79257 | AAY79257 Putative |
| 20 | 31 | 100.0 | 144 | 4 AAU00478 | AAU00478 Rat INTBR |
| 21 | 31 | 100.0 | 144 | 6 ABO32630 | ABO32630 Secreted |
| 22 | 31 | 100.0 | 147 | 3 AAY79267 | AAY79267 Mouse put |
| 23 | 31 | 100.0 | 174 | 4 AAU00477 | AAU00477 Murine IN |
| 24 | 31 | 100.0 | 174 | 6 ABO32622 | ABO32622 Secreted |
| 25 | 31 | 100.0 | 179 | 6 AAB37329 | AAB37329 Human V1C |

| | | | | | |
|----|----|-------|-----|------------|--------------------|
| 26 | 31 | 100.0 | 211 | 2 AAW29676 | AAW29676 Homo sapi |
| 27 | 31 | 100.0 | 219 | 6 ABO32609 | ABO32609 Secreted |
| 28 | 31 | 100.0 | 239 | 2 AAR43679 | AAR43679 Single ch |
| 29 | 31 | 100.0 | 239 | 2 AAR99649 | AAR99649 Single ch |
| 30 | 31 | 100.0 | 239 | 2 AAW02191 | AAW02191 18-2-3/TR |
| 31 | 31 | 100.0 | 240 | 2 AAW05133 | AAW05133 Single ch |
| 32 | 31 | 100.0 | 240 | 4 AAM93344 | AAM93344 Human pol |
| 33 | 31 | 100.0 | 241 | 2 AAR06482 | AAR06482 18-2-3-/T |
| 34 | 31 | 100.0 | 242 | 2 AAR06483 | AAR06483 18-2-3-/T |
| 35 | 31 | 100.0 | 242 | 2 AAR43680 | AAR43680 Single ch |
| 36 | 31 | 100.0 | 242 | 2 AAR99650 | AAR99650 Single ch |
| 37 | 31 | 100.0 | 242 | 2 AAW02192 | AAW02192 18-2-3/TR |
| 38 | 31 | 100.0 | 251 | 5 ABP44994 | ABP44994 Human Bly |
| 39 | 31 | 100.0 | 265 | 3 AAY79266 | AAY79266 Human elo |
| 40 | 31 | 100.0 | 265 | 3 AAB29648 | AAB29648 Human mem |
| 41 | 31 | 100.0 | 265 | 4 AAU00476 | AAU00476 Human INT |
| 42 | 31 | 100.0 | 265 | 5 AAU87832 | AAU87832 Human elo |
| 43 | 31 | 100.0 | 265 | 5 AAE24555 | AAE24555 Human ELG |
| 44 | 31 | 100.0 | 265 | 6 AAG79838 | AAG79838 ADSL rela |
| 45 | 31 | 100.0 | 265 | 6 ABO32607 | ABO32607 Secreted |

ALIGNMENTS

RESULT 1
AAW08938
ID AAW08938 standard; peptide; 5 AA.
XX
AC AAW08938;
DT 18-SEP-1997 (first entry)
DE CDR1 from heavy chain variable region of reshaped H225 antibody.
XX
KW Complementarity determining region 1; CDR1; heavy chain; variable region;
KW reshaped; human; monoclonal; antibody; H225; epidermal growth factor;
KW EGF; receptor; inhibition; growth; tumour; cell; late stage; prostatic;
KW prostate.
XX
OS Homo sapiens.
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
XX
PS 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
XX
PB (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
PT WPI; 1997-051897/05.
XX
PS Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PT Claim 1; Page 83; 112pp; English.
XX
PS The present peptide is the complementarity determining region 1 (CDR1)
XX from the heavy chain variable region of the reshaped human monoclonal
XX antibody (Mab) H225, which is specific for the human epidermal growth
XX factor (EGF) receptor. The Mab, or a fragment, can be used to inhibit the
XX growth of tumour cells, especially late stage prostatic tumour cells in
XX humans, optionally conjugated to a cytotoxic agent, especially
XX doxorubicin, taxol or cisplatin, or a signal transduction, ras or cell
XX cycle inhibitor
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 31; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 1 NYGVH 5

RESULT 2
 AAY59311
 ID AAY59311 standard; peptide; 5 AA.
 AC AAY59311;
 XX
 DT 07-MAR-2000 (first entry)
 DE Heavy chain hypervariable region, CDR1.
 XX
 KW Hypervariable region; complementarity determining region; CDR; tumour;
 KW single chain antibody; growth inhibitor; human; tumourigenesis; therapy;
 KW protein receptor tyrosine kinase; heavy chain.
 XX
 OS Mus sp.
 XX
 PN WO9960023-A1.
 XX
 PD 25-NOV-1999.
 XX
 PE 14-MAY-1999; 99WO-US010741.
 XX
 PR 15-MAY-1998; 98US-00079612.
 PR 15-MAY-1998; 98US-0085613P.
 PR 07-DEC-1998; 98US-00206139.
 XX
 PA (IMCL-) IMCLONE SYSTEMS INC.
 PA (UABR-) UAB RES FOUND.
 XX
 PI Waksal HW, Saleh MN, Robert F, Buchsbaum DJ;
 XX
 DR WPI; 2000-062440/05.
 DR N-PSDB; AA248625.
 XX
 PT Treatment of human tumors, using a combination of radiation and a non-
 PT radiolabeled protein receptor tyrosine kinase inhibitor.
 XX
 PS Disclosure; Page 14; 31pp; English.
 XX
 CC This sequence is the hypervariable region CDR1 (complementarity
 CC determining region 1) of the heavy chain of a single chain antibody
 CC derived from the murine antibody 225. The invention relates to a method
 CC for inhibiting the growth of tumours in human patients by treating with
 CC an effective amount of a combination of radiation and a non-radiolabelled
 CC protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
 CC which can lead to tumourigenesis. The method can be used in the treatment
 CC of tumours of e.g. breast, lung, colon, kidney, bladder, head and neck,
 CC ovary, prostate or brain. The administration of a suitable antibody to
 CC the patient makes the tumour more susceptible to radiotherapy
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 31; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 1 NYGVH 5

RESULT 3
 AAB37952
 ID AAB37952 standard; protein; 5 AA.
 AC AAB37952;
 XX
 DT 12-MAR-2001 (first entry)
 XX
 DE Anti-EGFR monoclonal antibody H chain V region CDR1 peptide sequence.
 XX
 KW Refractory tumour growth inhibition; epidermal growth factor receptor;
 KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
 KW complementarity determining region; CDR.
 XX
 OS Mus sp.
 XX
 PN WO200069459-A1.
 XX
 PD 23-NOV-2000.
 XX
 PE 01-MAY-2000; 2000WO-US011756.
 XX
 PR 14-MAY-1999; 99US-00312284.
 PR 13-AUG-1999; 99US-00374028.
 XX
 PA (IMCL-) IMCLONE SYSTEMS INC.
 XX
 PI Waksal HW;
 XX
 DR WPI; 2001-016160/02.
 DR N-PSDB; AAC83235.
 XX
 PT Epidermal growth factor receptor/human epidermal growth factor receptor-1
 PT antagonist for inhibiting the growth of refractory tumors.
 XX
 PS Disclosure; Page 13; 31pp; English.
 XX
 CC This invention relates to a method for inhibiting the growth of
 CC refractory tumors that are stimulated by a ligand of epidermal growth
 CC factor receptor (EGFR) in human patients. The method involves treating
 CC the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
 CC optionally with a chemotherapeutic agent or radiation. The antagonist can
 CC be for example a chimeric anti-EGFR monoclonal antibody, C225. The
 CC EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
 CC tumors such as tumours of breast, heart, lung, small intestine, colon,
 CC spleen, kidney, bladder, head and neck, ovary, prostate, cervix and
 CC skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
 CC liver, preferably squamous cell carcinomas. The present sequence
 CC represents the heavy chain variable region complementarity determining
 CC region 1 amino acid sequence of the chimeric anti-EGFR monoclonal
 CC antibody C225 which is used in an example illustrating the method of the
 CC invention
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 31; DB 4; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 1 NYGVH 5

RESULT 4
 AAU77784
 ID AAU77784 standard; peptide; 5 AA.
 XX
 AC AAU77784;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Mouse heavy chain hypervariable region (CDR1) of 225 antibody.
 XX
 KW Mouse; heavy chain; antibody; hyperproliferative disease;

KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
 KW seborrheic keratosis; warts; keloid scars; eczema; 255 antibody;
 KW hypervariable region; CDRI; EGFR inhibitor.
 XX
 OS Mus sp.
 XX
 XX WO200211677-A2.
 PN
 XX
 PD 14-FEB-2002.
 XX
 XX 09-AUG-2001; 2001WO-US041647.
 PF
 XX 09-AUG-2000; 2000US-00635974.
 PR
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA
 XX Teufel T;
 PI
 XX WPI; 2002-257423/30.
 DR
 XX N-PSDB; ABK11440.
 XX
 XX Treating a mammal with hyperproliferative disease especially psoriasis,
 PT stimulated by ligand of member of epidermal growth factor family of
 PT receptors, by administering antagonist of the receptor.
 XX
 XX Disclosure; Page 11; 28pp; English.
 PS
 XX This invention relates to a novel method for treating a mammal with
 XX hyperproliferative disease stimulated by a ligand of a member of the
 CC epidermal growth factor (EGF) family of receptors. The method involves
 CC administering an antibody or a defective receptor that is an antagonist
 CC of a member of the EGF receptor family, or a combination of the
 CC antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
 CC The antibody used in the method of the invention acts as an epidermal
 CC growth factor receptor (EGFR) antagonist by inhibiting EGFR /HER1
 CC phosphorylation. The method of the invention is useful for treating a
 CC mammal with hyperproliferative disease such as psoriasis, actinic
 CC keratoses, seborrheic keratoses, warts, keloid scars and eczema
 CC stimulated by a ligand of a member of the EGF family of receptor. This
 CC sequence represents the murine anti-EGF 255 antibody heavy chain (VH)
 CC hypervariable region (CDRI) used as an inhibitor of EGFR in the method of
 CC the invention
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 31; DB 5; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 DB |||||
 1 NYGVH 5
 RESULT 5
 ABO32615
 ID ABO32615 standard; peptide; 19 AA.
 XX
 AC ABO32615;
 XX
 DT 17-SEP-2003 (first entry)
 XX
 XX Secreted polypeptide-related peptide #42.
 DE
 XX Human; TANGO; INTERCEPT; secreted polypeptide; immune disorder;
 KW hormonal disorder; proliferative disorder; cancer; thyroid disorder;
 KW diabetes; multiple sclerosis; lupus; neurological disorder; anaemia;
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
 KW myocardial infarction; congestive heart disease; blood platelet disorder;
 KW thrombocytopenia; blood vessel; atherosclerosis; vasculitis.
 XX
 XX Homo sapiens.
 XX

PN US2003022279-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 12-JAN-2001; 2001US-00759130.
 XX
 XX 14-JUN-1999; 99US-00333159.
 PR 29-JUN-1999; 99US-00342364.
 PR 10-SEP-1999; 99US-00393996.
 PR 19-OCT-1999; 99US-00420707.
 PR 07-JAN-2000; 2000US-00479249.
 PR 27-APR-2000; 2000US-00559497.
 PR 24-MAY-2000; 2000US-00578063.
 PR 16-JUN-2000; 2000US-00596194.
 PR 23-JUN-2000; 2000US-00602871.
 PR 30-JUN-2000; 2000US-00608452.
 XX (FRAS/) FRASER C C.
 PA (BARN/) BARNES T M.
 PA (SHAR/) SHARP J D.
 PA (KIRS/) KIRST S J.
 PA (MYER/) MYERS P S.
 PA (LEIB/) LEIBY K R.
 PA (HOLT/) HOLTZMAN D A.
 PA (MCCA/) MCCARTHY S A.
 PA (WRIG/) WRIGHTON N.
 PA (MACK/) MACKAY C R.
 PA (GOOD/) GOODEARL A D J.
 XX
 XX Fraser CC, Barnes TM, Sharp JD, Kirst SJ, Myers PS, Leiby KR;
 PI Holtzman DA, McCarthy SA, Wrighton N, Mackay CR, Goodearl ADJ;
 PI WPI; 2003-456290/43.
 XX
 XX New nucleic acid molecule encoding a secreted protein (e.g. TANGO 202,
 PT TANGO 210 or INTERCEPT 217), useful for diagnosing, preventing or
 PT treating disorders such as cancer, diabetes or atherosclerosis, and in
 PT forensic biology.
 XX
 XX Claim 9; SEQ ID NO 231; 482pp; English.
 XX
 XX The invention relates to secreted polypeptide-related proteins and
 CC nucleic acids (TANGO and INTERCEPT proteins and nucleic acids). The
 CC nucleic acids, proteins and antibodies specific to the proteins are
 CC useful in screening assays, predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, monitoring clinical trials and pharmacogenetics) and
 CC prophylactic and therapeutic methods. The sequences are used in
 CC diagnosing, preventing or treating proliferative disorders (e.g.
 CC cancer), hormonal disorders (e.g. diabetes or thyroid disorders), immune
 CC disorders (e.g. multiple sclerosis or lupus), neurological disorders
 CC (e.g. Alzheimer's disease or Parkinson's disease), cardiovascular
 CC disorders (e.g. myocardial infarction or congestive heart disease), blood
 CC platelet disorders (e.g. thrombocytopenia or anaemia) and disorders
 CC involving blood vessels (e.g. atherosclerosis or vasculitis). The nucleic
 CC acids may also be used in chromosome mapping, tissue typing and forensic
 CC biology, and as surrogate markers. This sequence represents a secreted
 CC polypeptide-related peptide of the invention. Note: The sequence data for
 CC this patent was obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 XX Sequence 19 AA;
 SQ
 Query Match 100.0%; Score 31; DB 6; Length 19;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 DB |||||
 6 NYGVH 10
 RESULT 6
 AAV93396

ID AAM93396 standard; protein; 107 AA.
 AC AAM93396;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 2992.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN BP1130094-A2.
 PD
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114089.
 XX
 PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94317.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 PS Claim 8; SEQ ID NO 2992; 1380pp + Sequence Listing; English.
 CC
 CC The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 CC format directly from EPO
 XX
 SQ Sequence 107 AA;
 Query Match 100.0%; Score 31; DB 4; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 DB 12 NYGVH 16
 |||||
 RESULT 7
 AAW05823
 ID AAW05823 standard; protein; 116 AA.
 XX
 AC AAW05823;
 XX
 DT 16-OCT-2003 (revised)
 DT 27-JAN-1997 (first entry)
 XX
 DE Humanised 1D10 antibody heavy chain variable region.
 XX
 KW B-cell lymphoma; humanised antibody; bispecific antibody; myeloma;
 KW leukaemia; hybridoma; monoclonal antibody.
 XX

OS Homo; sapiens.
 OS Mus sp.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 27 /note= "human framework residue 27 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 29 /note= "human framework residue 29 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 30 /note= "human framework residue 30 is substd. by
 FT equivalent murine framework residue"
 FT Region 31..35 /label= CDR1
 FT Misc-difference 37 /note= "human framework residue 37 is substd. by
 FT equivalent murine framework residue"
 FT Region 50..65 /label= CDR2
 FT Misc-difference 67 /note= "human framework residue 68 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 71 /note= "human framework residue 71 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 73 /note= "human framework residue 73 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 78 /note= "human framework residue 78 is substd. by
 FT equivalent murine framework residue"
 FT Misc-difference 82 /note= "human framework residue 82 is substd. by a
 FT consensus amino acid of the equivalent murine framework
 FT residue"
 FT Region 98..105 /label= CDR3
 XX WO9626964-A1.
 PD 06-SEP-1996.
 XX
 PF 29-FEB-1996; 96WO-US0002754.
 XX
 PR 01-MAR-1995; 95US-00397411.
 XX
 PA (PROT-) PROTEIN DESIGN LABS INC.
 PA (IOWA-) IOWA IMMUNOTHERAPY INVESTIGATORS.
 XX
 PI Weiner G, Gingrich R, Link BK, Tso JY;
 XX WPI; 1996-412742/41.
 DR
 PT New bi-specific antibody reactive with both T or NK cells and malignant B
 PT cells - also their humanised forms and hybridomas producing them, useful
 PT for treating or preventing leukaemia, lymphoma and myeloma.
 XX
 PS Claim 14; Fig 4b; 85pp; English.
 XX
 CC The variable region (AAW05823) of the humanised 1D10 antibody heavy chain
 CC (AAW05829) consists of human IC4 heavy chain variable region framework
 CC (substd. at 9 positions with mouse or consensus human amino acids) and
 CC complementarity determining regions from the murine 1D10 antibody
 CC specific for a 28/32 kDa heterodimeric antigen present on the surface of
 CC malignant B-cells. It can be coexpressed with humanised 1D10 light chain
 CC (see also AAW05828) in e.g. mouse myeloma NSO cells. Humanised antibody
 CC fragments can be incorporated into novel bispecific antibodies reactive
 CC with both effector cell antigens (see also AAW05824-25, AAW05826 and
 CC AAW05830) and malignant B-cells. (Updated on 16-OCT-2003 to standardise
 CC OS field)
 XX

```
SQ      Sequence 116 AA;
Query Match      100.0%; Score 31; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NYGVH 5
Db      31 NYGVH 35

RESULT 8
AAG33864
ID      AAG33864 standard; protein; 117 AA.
XX
AC      AAG33864;
XX
DT      19-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 41114.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000BP-00301439.
XX
PR      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0123160P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      16-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      17-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142350P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146388P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
PR      05-AUG-1999; 99US-0147260P.
PR      06-AUG-1999; 99US-0147303P.
PR      06-AUG-1999; 99US-0147416P.
PR      09-AUG-1999; 99US-0147493P.
```



```

Db      31 NYGVH 35
|||||
RESULT 10
AAW08950
ID AAW08950 standard; protein; 119 AA.
XX
AC AAW08950;
XX
DT 18-SEP-1997 (first entry)
XX
DE Heavy chain variable region of 225RA antibody.
XX
KW Heavy chain; reshaped; monoclonal; antibody; 225RA; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
PI MPI; 1997-051897/05.
XX
DR Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PS Claim 30; Fig 22; 112pp; English.
XX
CC The present sequence is the heavy chain variable region of the reshaped
CC human monoclonal antibody (Mab) H225, 225RHA. The Mab is specific for the
CC human epidermal growth factor (EGF) receptor. The Mab, or a fragment, can
CC be used to inhibit the growth of tumour cells, especially late stage
CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
CC agent, especially doxorubicin, taxol or cisplatin, or a signal
CC transduction, ras or cell cycle inhibitor
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 31; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYGVH 5

```

Query Match 100.0%; Score 31; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYGVH 5

```

Db          31 NYGVH 35
|||||
RESULT 12
AAW08951
ID AAW08951 standard; protein; 119 AA.
XX
XX AC AAW08951;
XX
XX DT 18-SEP-1997 (first entry)
XX
XX DE Heavy chain variable region of 22SRB antibody.
XX
XX KW Heavy chain; reshaped; monoclonal; antibody; 22SRB; human;
XX KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX KW late stage; prostatic; prostate; variable region; framework;
XX KW complementarity determining region; CDR.
XX
XX OS Homo sapiens.
XX
XX FH Key
XX FT Region 1..30
XX FT /label= framework_1
XX FT Region 31..35
XX FT /label= CDR_1
XX FT Region 36..49
XX FT /label= framework_2
XX FT Region 50..65
XX FT /label= CDR_2
XX FT Region 66..97
XX FT /label= framework_3
XX FT Region 98..108
XX FT /label= CDR_3
XX FT Region 109..119
XX FT /label= framework_4
XX
XX PN WO9640210-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 07-JUN-1996; 96WO-US009847.
XX
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (MRCC-) MRC COLLABORATIVE CENT.
XX
XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
XX PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX PS Claim 30; Fig 22; 112pp; English.
XX
XX CC The present sequence is the heavy chain variable region of the reshaped
XX CC human monoclonal antibody (MAB) H25, 22SRB. The MAB is specific for the
XX CC human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX CC be used to inhibit the growth of tumour cells, especially late stage
XX CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX CC agent, especially doxorubicin, taxol or cisplatin, or a signal
XX CC transduction, ras or cell cycle inhibitor
XX
XX SQ Sequence 119 AA;

Query Match 100.0%; Score 31; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db          31 NYGVH 35
|||||
RESULT 13
AAW08954
ID AAW08954 standard; protein; 119 AA.
XX
XX AC AAW08954;
XX
XX DT 18-SEP-1997 (first entry)
XX
XX DE Heavy chain variable region of 22SRE antibody.
XX
XX KW Heavy chain; reshaped; monoclonal; antibody; 22SRE; human;
XX KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX KW late stage; prostatic; prostate; variable region; framework;
XX KW complementarity determining region; CDR.
XX
XX OS Homo sapiens.
XX
XX FH Key
XX FT Region 1..30
XX FT /label= framework_1
XX FT Region 31..35
XX FT /label= CDR_1
XX FT Region 36..49
XX FT /label= framework_2
XX FT Region 50..65
XX FT /label= CDR_2
XX FT Region 66..97
XX FT /label= framework_3
XX FT Region 98..108
XX FT /label= CDR_3
XX FT Region 109..119
XX FT /label= framework_4
XX
XX PN WO9640210-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 07-JUN-1996; 96WO-US009847.
XX
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (MRCC-) MRC COLLABORATIVE CENT.
XX
XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
XX PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX PS Claim 30; Fig 22; 112pp; English.
XX
XX CC The present sequence is the heavy chain variable region of the reshaped
XX CC human monoclonal antibody (MAB) H225, 22SRE. The MAB is specific for the
XX CC human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX CC be used to inhibit the growth of tumour cells, especially late stage
XX CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX CC agent, especially doxorubicin, taxol or cisplatin, or a signal
XX CC transduction, ras or cell cycle inhibitor
XX
XX SQ Sequence 119 AA;

Query Match 100.0%; Score 31; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

```

```

Db          31 NYGVH 35
|||||
RESULT 14
AA90818
ID AA90818 standard; protein; 119 AA.
XX
XX AC AA90818;
XX
XX DT 29-AUG-2000 (first entry)
XX
XX DE 260F9 hybridoma VL domain SEQ ID NO:14.
XX
XX KW Antigen binding site; immunoglobulin; cancer antigen; immunological;
XX antibody; tumour; human; mucin; cancer; cytostatic; hybridoma;
XX specific binding assay; affinity purification; drug targeting;
XX KW toxin targeting; imaging; genetic; therapeutic.
XX
XX OS Homo sapiens.
XX
XX PN US6054561-A.
XX
XX PD 25-APR-2000.
XX
XX PF 07-JUN-1995; 95US-00483749.
XX
XX PR 08-FEB-1984; 84US-00577976.
XX
XX PR 11-JAN-1985; 85US-00690750.
XX
XX PR 21-MAR-1986; 86US-00842476.
XX
XX PR 08-MAY-1988; 88US-00190778.
XX
XX PR 11-AUG-1994; 94US-00289881.
XX
XX (CHIR ) CHIRON CORP.
XX
XX PI Ring DB;
XX
XX DR WPI; 2000-338508/29.
XX
XX DR N-PSDB; AAA38902.
XX
XX PT Monoclonal antibody capable of binding to human breast cancer antigen
XX useful for affinity purification, drug or toxin targeting, imaging, and
XX treating cancer.
XX
XX PS Disclosure; Fig 7; 57pp; English.
XX
XX CC The present invention describes a monoclonal antibody (Mab) (I) that
XX binds to a human breast cancer antigen that is also bound by Mab 454C11
XX and 520C9 (produced hybridoma ATCC HB8484 and HB8696, respectively). Also
XX described is a hybridoma that produces (I). (I) is useful in specific
XX binding assays, affinity purification, drug or toxin targeting, imaging,
XX CC and genetic or immunological therapeutics for various cancers. The
XX present sequence represents a VH domain derived from a 260F9 hybridoma,
XX which is used in the exemplification of the present invention
XX
XX SQ Sequence 119 AA;
XX
XX Query Match 100.0%; Score 31; DB 3; Length 119;
XX Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NYGVH 5
XX |||||
XX Db 31 NYGVH 35
XX
XX RESULT 15
XX AAB00210
XX ID AAB00210 standard; protein; 130 AA.
XX
XX AC AAB00210;
XX
XX DT 08-FEB-2001 (first entry)
XX
XX DE Human sequence (BAC207d4) related to PUFA elongase.
XX
XX KW Elongase; polyunsaturated fatty acid; PUFA; dietary supplement;
XX pharmaceutical; foodstuff; food; di-homo-gamma-linoleic acid;
XX arachidonic acid; eicosapentaenoic acid; docosatrienoic acid;
XX KW docosatetraenoic acid; docosapentaenoic acid; docosahexaenoic acid;
XX human.
XX
XX OS Homo sapiens.
XX
XX PN WO200055330-A1.
XX
XX PD 21-SEP-2000.
XX
XX XX 20-MAR-2000; 2000WO-GB001035.
XX
XX PR 18-MAR-1999; 99GB-00006307.
XX
XX PR 18-FEB-2000; 2000GB-00003869.
XX
XX (UYBR-) UNIV BRISTOL.
XX
XX PI Napier JA;
XX
XX DR WPI; 2000-647074/62.
XX
XX PT Novel isolated polypeptide comprising functional long chain
XX polyunsaturated fatty acid (PUFA) elongase of Caenorhabditis elegans used
XX to produce PUFA for foodstuff, dietary supplement or pharmaceutical
XX composition.
XX
XX PS Disclosure; Fig 10; 42pp; English.
XX
XX CC New elongase polypeptides which are functional long chain polyunsaturated
XX fatty acid (PUFA) elongase polypeptides are described. The elongase
XX catalyses an elongase reaction to produce PUFA (a 24 carbon fatty acid
XX with at least 4 double bonds) such as di-homo-gamma-linoleic acid
XX CC (20:3Delta8,11,14), arachidonic acid (20:4Delta5,8,11,14),
XX eicosapentaenoic acid (20:5Delta5,8,11,14,17), docosatrienoic acid
XX CC (22:3Delta3,16,19), docosatetraenoic acid (22:4Delta7,10,13,16),
XX docosapentaenoic acid (22:5Delta7,10,13,16,19) or docosahexaenoic acid
XX CC (22:6Delta4,7,10,13,16,19). PUFA produced is useful in foodstuffs,
XX dietary supplements and in pharmaceutical compositions which can then be
XX used to elavate PUFA levels of an animal or plant
XX
XX SQ Sequence 130 AA;
XX
XX Query Match 100.0%; Score 31; DB 3; Length 130;
XX Best Local Similarity 100.0%; Pred. NO. 1.3e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NYGVH 5
XX |||||
XX Db 45 NYGVH 49
XX
XX Search completed: October 6, 2004, 16:29:42
XX Job time : 29.2632 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:13 ; Search time 16.7544 Seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-2

Perfect score: 31

Sequence: 1 NYGVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTRMBL_25:*
- 2: sp_archaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_phase:*
- 11: sp_plant:*
- 12: sp_rodent:*
- 13: sp_virus:*
- 14: sp_vertebrate:*
- 15: sp_unclassified:*
- 16: sp_rvirus:*
- 17: sp_bacteriap:*
- 18: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 31 | 100.0 | 240 | Q8NCD1 | Q8ncd1 homo sapien |
| 2 | 31 | 100.0 | 255 | Q8TEP9 | Q8tep9 homo sapien |
| 3 | 31 | 100.0 | 263 | Q8YMI3 | Q8ym13 anabaena sp |
| 4 | 31 | 100.0 | 265 | Q8H5J4 | Q8h5j4 homo sapien |
| 5 | 31 | 100.0 | 267 | Q8ZOL6 | Q8zol6 rattus norv |
| 6 | 31 | 100.0 | 267 | Q8CE45 | Q8ce45 mus musculu |
| 7 | 31 | 100.0 | 267 | Q9ZOL5 | Q9zol5 mus musculu |
| 8 | 31 | 100.0 | 338 | Q9XX32 | Q9xx32 caenorhabdi |
| 9 | 31 | 100.0 | 371 | Q93I29 | Q93i29 streptomyce |
| 10 | 31 | 100.0 | 400 | Q9BTV4 | Q9btv4 homo sapien |
| 11 | 31 | 100.0 | 400 | Q8NC30 | Q8nc30 homo sapien |
| 12 | 31 | 100.0 | 400 | Q8EF19 | Q8ef19 homo sapien |
| 13 | 31 | 100.0 | 400 | Q9H076 | Q9h076 homo sapien |
| 14 | 31 | 100.0 | 400 | Q9DBS1 | Q9db1 mus musculu |
| 15 | 31 | 100.0 | 403 | Q8PKV6 | Q8pkv6 xanthomonas |
| 16 | 31 | 100.0 | 418 | Q8EBW6 | Q8ebw6 shewanella |

| | | | | | |
|----|----|-------|------|----|--------|
| 17 | 31 | 100.0 | 475 | 16 | P74540 |
| 18 | 31 | 100.0 | 482 | 10 | Q9MA74 |
| 19 | 31 | 100.0 | 490 | 10 | Q9SXL2 |
| 20 | 31 | 100.0 | 570 | 5 | Q9GYB6 |
| 21 | 31 | 100.0 | 622 | 16 | Q8DGA1 |
| 22 | 31 | 100.0 | 628 | 5 | Q8EAS5 |
| 23 | 31 | 100.0 | 655 | 9 | Q8HAK0 |
| 24 | 31 | 100.0 | 796 | 16 | Q8E862 |
| 25 | 31 | 100.0 | 856 | 5 | Q8E833 |
| 26 | 31 | 100.0 | 887 | 5 | Q8E8N8 |
| 27 | 31 | 100.0 | 887 | 5 | Q8E8G4 |
| 28 | 31 | 100.0 | 900 | 16 | Q8FSD7 |
| 29 | 31 | 100.0 | 913 | 5 | Q8E8D7 |
| 30 | 31 | 100.0 | 922 | 5 | Q8E8U5 |
| 31 | 31 | 100.0 | 959 | 10 | Q8SAT5 |
| 32 | 31 | 100.0 | 1011 | 2 | Q8L2B5 |
| 33 | 31 | 100.0 | 1263 | 5 | Q8TOM2 |
| 34 | 31 | 100.0 | 1439 | 5 | Q8TOM2 |
| 35 | 31 | 100.0 | 1462 | 5 | Q8V5Q8 |
| 36 | 31 | 100.0 | 1465 | 4 | Q8Z3Y2 |
| 37 | 30 | 96.8 | 169 | 16 | Q8XK41 |
| 38 | 30 | 96.8 | 170 | 16 | Q8YNO3 |
| 39 | 30 | 96.8 | 238 | 17 | Q8ZYH1 |
| 40 | 30 | 96.8 | 347 | 2 | Q8WZY5 |
| 41 | 30 | 96.8 | 349 | 3 | Q871K7 |
| 42 | 30 | 96.8 | 356 | 16 | Q8E4G8 |
| 43 | 30 | 96.8 | 356 | 16 | Q8DYW0 |
| 44 | 30 | 96.8 | 368 | 16 | Q8Y164 |
| 45 | 30 | 96.8 | 391 | 17 | Q8TU05 |

ALIGNMENTS

RESULT 1

Q8NCD1 PRELIMINARY; PRT; 240 AA.
AC Q8NCD1;
AT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein FLJ90332.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Oka T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Oka T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074813; BAC1125.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR002076; GNS1_SUR4.
DR Pfam; PF01151; ELO; 1.
DR PROSITE; PS01188; ELO; 1.
KW Hypothetical protein.
SQ SEQUENCE 240 AA; 28271 MW; 39A101DCF141818E CRC64;

Query Match 100.0%; Score 31; DB 4; Length 240;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYGVH 5

Db 145 NYGVH 149

RESULT 2

Q8TEP9 PRELIMINARY; PRT; 255 AA.
 AC Q8TEP9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE FLJ00144 protein (Fragment).
 GN FLJ00144.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human
 spleen.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074073; BAB84899.1; -;
 FT NON TER 1
 SQ SEQUENCE 255 AA; 28454 MW; BF65941BB2012999 CRC64;
 Query Match 100.0%; Score 31; DB 4; Length 255;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 35 NYGVH 39
 RESULT 3
 Q8YM13 PRELIMINARY; PRT; 263 AA.
 AC Q8YM13;
 DT 01-VAR-2002 (TrEMBLrel. 20, Created)
 DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Hypothetical protein A114950.
 GN A114950.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda N., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003598; BAB76649.1; -;
 DR FIR; AF2424; AF2424.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 263 AA; 31349 MW; DF8866BDD09F6DCF CRC64;
 Query Match 100.0%; Score 31; DB 16; Length 263;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 145 NYGVH 149

RESULT 4
 Q9H5J4 PRELIMINARY; PRT; 265 AA.
 ID Q9H5J4
 AC Q9H5J4;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ23378.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027031; BAB15632.1; -;
 DR EMBL; BC001305; AA01305.1; -;
 DR Genew; HGNC:15829; ELQVL6.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR002076; GNS1_SUR4.
 DR Pfam; PF01151; ELO; 1.
 DR PROSITE; PS01188; ELO; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 265 AA; 31376 MW; 01234E0E6F6CE341 CRC64;
 Query Match 100.0%; Score 31; DB 4; Length 265;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 170 NYGVH 174
 RESULT 5
 Q920L6 PRELIMINARY; PRT; 267 AA.
 ID Q920L6;
 AC Q920L6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Fatty acid elongase 2.
 GN RELO2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague Dawley; TISSUE=Liver;
 RA Inagaki K., Aki T., Fukuda Y., Kawamoto S., Shigeta S., Ono K.,
 RA Suzuki O.;
 RT "Identification and expression of a mammalian fatty acid elongase
 RT involved in de novo biosynthesis of C18 fatty acids.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB071986; BAB69888.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR002076; GNS1_SUR4.
 DR Pfam; PF01151; ELO; 1.
 DR PROSITE; PS01188; ELO; 1.
 SQ SEQUENCE 267 AA; 31624 MW; 4026C9CB33ED0743 CRC64;
 Query Match 100.0%; Score 31; DB 11; Length 267;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 111

Db 170 NYGVH 174

RESULT 6

Q8CE45 PRELIMINARY; PRT; 267 AA.

AC Q8CE45; 2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Long chain fatty acyl elongase.

GN ELOVL6 OR LCE

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

EX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR ENBL; AK029029; BAC26252.1; -.

DR MGD; MGI:2156528; Elov16.

DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.

DR GO; GO:0016747; P:transferase activity, transferring groups o...; IDA.

DR GO; GO:0030497; P:fatty acid elongation; IDA.

DR InterPro; IPR002076; GNS1_SUR4.

DR Pfam; PF01151; ELO; 1.

SQ SEQUENCE 267 AA; 31601 MW; 4031F9C503E4DD23 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 267;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 170 NYGVH 174

RESULT 7

Q920L5 PRELIMINARY; PRT; 267 AA.

AC Q920L5; 2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Fatty acyl elongase (Long-chain fatty-acyl elongase) (Myelination

DE associated SUR4-like protein).

GN ELOVL6 OR FAE OR LCE OR MASR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Liver;

RA Matsuzaka T.; Shimano H.;

RA "murine complete cds for new fatty acyl elongase similar to cig30

RT (U97107).";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

EX MEDLINE=21576178; PubMed=11567032;

RA Moon Y.A.; Shah N.A.; Mohapatra S.; Warrington J.A.; Horton J.D.;

RA "Identification of a Mammalian Long Chain Fatty Acyl Elongase

RT Regulated by Sterol Regulatory Element-binding Proteins.";

RL J. Biol. Chem. 276:45358-45366 (2001).

RN [3]

RP SEQUENCE FROM N.A.

RA Nagarajan R., Le N.H., Mahoney H., Araki T., Milbrandt J.D.;

RT "Deciphering Peripheral Nerve Myelination Using Schwann Cell

RT Expression Profiling.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=CZECH II; TISSUE=Breast tumor;

RA Strausberg R.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR ENBL; AB072039; BAB68544.1; -.

DR ENBL; AV053453; AAL14239.1; -.

DR ENBL; AF480860; AAM13450.1; -.

DR ENBL; BC051041; AHS1041.1; -.

DR MGD; MGI:2156528; Elov16.

DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.

DR GO; GO:0016747; P:transferase activity, transferring groups o...; IDA.

DR GO; GO:0030497; P:fatty acid elongation; IDA.

DR InterPro; IPR002076; GNS1_SUR4.

DR Pfam; PF01151; ELO; 1.

DR PROSITE; PS01188; ELO; 1.

SQ SEQUENCE 267 AA; 31610 MW; 4026C9CB33FDD23 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 267;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 170 NYGVH 174

RESULT 8

Q9XX32 PRELIMINARY; PRT; 338 AA.

ID Q9XX32; 1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Y32B12B.6 protein.

GN Y32B12B.6.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Gardner A.E.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology.";

RL Science 282:2012-2018 (1998).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

DR ENBL; AL031632; CAA31008.1; -.

DR PIR; T26588; T26588.

DR HSP; P20393; I86Y.

DR Wormpep; Y32B12B.6; CEI9086.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.

DR GO; GO:0003700; P:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR GO; GO:0006350; P:transcription; IEA.

DR InterPro; IPR000536; Hormone_rec_lig.

DR InterPro; IPR009546; Str_ncl_receptor.

DR InterPro; IPR001628; Znf_C4steroid.

DR Pfam; PF00104; hormone_rec; 1.

DR Pfam; PF00105; zf-C4; 1.

DR PRINTS; PR00047; STROIDFINGER.

DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger;
 SQ SEQUENCE 338 AA; 39146 MW; 76526E469D05868A CRC64;
 Query Match 100.0%; Score 31; DB 5; Length 338;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NYGVH 5
 Db 26 NYGVH 30
 RESULT 9
 Q93129 PRELIMINARY; PRT; 371 AA.
 AC Q93129; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Putative secreted peptidase.
 GN SC07521 OR SCBAC25F8.13C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939131; CAC42148.1; -.
 DR HSSP; P15555; ICEP.
 DR InterPro; IPR001486; Beta_lactamase.
 DR Pfam; PF00144; beta-lactamase; 1.
 KW Complete proteome.
 SQ SEQUENCE 371 AA; 39201 MW; 5A54095CBFC8C27 CRC64;
 Query Match 100.0%; Score 31; DB 16; Length 371;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NYGVH 5
 Db 226 NYGVH 230
 RESULT 10
 Q9BTV4 PRELIMINARY; PRT; 400 AA.
 AC Q9BTV4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FLJ14971 (Hypothetical protein FLJ14851).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid, and Placenta;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003125; AAH03125.1; -.
 DR EMBL; AK027877; BAB55425.1; -.
 DR EMBL; AK027757; BAB55348.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 400 AA; 44875 MW; 70FDD44ED1AA11DF CRC64;
 Query Match 100.0%; Score 31; DB 4; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NYGVH 5
 Db 105 NYGVH 109
 RESULT 11
 Q8NC30 PRELIMINARY; PRT; 400 AA.
 AC Q8NC30; (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DE Hypothetical protein FLJ90529.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuho Y., Ota T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK075010; BAC11350.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 400 AA; 44859 MW; 14F5401ADF9C32A6 CRC64;
 Query Match 100.0%; Score 31; DB 4; Length 400;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 NYGVH 5
 Db 105 NYGVH 109
 RESULT 12
 Q96F19 PRELIMINARY; PRT; 400 AA.
 AC Q96F19;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011719; AAH11719.1; -
KW Hypothetical protein.
SQ SEQUENCE 400 AA; 44847 MW; BD89DB4BB7113CF8 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 105 NYGVH 109

RESULT 13
Q9H076 PRELIMINARY; PRT; 400 AA.
ID Q9H076
AC Q9H076
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN DXFZP586G1919.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136916; CAB66850.1; -
KW Hypothetical protein.
SQ SEQUENCE 400 AA; 44831 MW; BD998B0BB73120E4 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 105 NYGVH 109

RESULT 14
Q9DBS1 PRELIMINARY; PRT; 400 AA.
ID Q9DBS1
AC Q9DBS1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1200015A22Rik protein (RIKEN cDNA 1200015A22 gene).
GN 1200015A22R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

```

```

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Caasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gusinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK004778; BAB23556.1; -
DR EMBL; BC024933; AAH24933.1; -
DR MGD; MGI:1921372; 1200015A22Rik.
SQ SEQUENCE 400 AA; 44783 MW; SA9732B4214D4316 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 105 NYGVH 109

RESULT 15
Q8PKV6 PRELIMINARY; PRT; 403 AA.
ID Q8PKV6
AC Q8PKV6
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Oxidoreductase.
GN XAC2051.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinoza L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AB011840; ABM36913.1; -
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006076; Fad_oxred.

```

DR Pfam; PF01266; DAO; 1.
KW Complete proteome.
SQ SEQUENCE 403 AA; 44543 MW; A8394D0E38880EEF CRC64;
Query Match 100.0%; Score 31; DB 16; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYGVH 5
Db 163 NYGVH 167

Search completed: October 6, 2004, 16:33:42
Job time : 22.7544 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 3.07018 Seconds
(without alignments)
84.800 Million cell updates/sec

Title: US-09-635-974A-2
Perfect score: 31
Sequence: 1 NYGVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 31 | 100.0 | 148 | 1 | P11734 locusta mag |
| 2 | 31 | 100.0 | 412 | 1 | PT56 YEAST |
| 3 | 31 | 100.0 | 435 | 1 | P49191 caenorhabdi |
| 4 | 31 | 100.0 | 1465 | 1 | MYM2 HUMAN |
| 5 | 31 | 100.0 | 6486 | 1 | TYC2 BREPA |
| 6 | 30 | 96.8 | 363 | 1 | CARA LEPIN |
| 7 | 30 | 96.8 | 449 | 1 | FUCU_BACSU |
| 8 | 30 | 96.8 | 581 | 1 | AMYI SCHPO |
| 9 | 30 | 96.8 | 680 | 1 | KALM HUMAN |
| 10 | 30 | 96.8 | 686 | 1 | FRE1 YEAST |
| 11 | 28 | 90.3 | 121 | 1 | HV3J HUMAN |
| 12 | 28 | 90.3 | 212 | 1 | NUIM HUMAN |
| 13 | 28 | 90.3 | 337 | 1 | NDFC_HUMAN |
| 14 | 28 | 90.3 | 337 | 1 | NDFC_MOUSE |
| 15 | 28 | 90.3 | 483 | 1 | TRKH ECOLI |
| 16 | 28 | 90.3 | 483 | 1 | TRKH SALTY |
| 17 | 28 | 90.3 | 659 | 1 | CAOP CAEL |
| 18 | 28 | 90.3 | 1207 | 1 | AT19 HUMAN |
| 19 | 28 | 90.3 | 1210 | 1 | AT19_MOUSE |
| 20 | 27 | 87.1 | 222 | 1 | HXCS_BRAE |
| 21 | 27 | 87.1 | 259 | 1 | ETFB_CLOAB |
| 22 | 27 | 87.1 | 270 | 1 | ELO3 HUMAN |
| 23 | 27 | 87.1 | 271 | 1 | ELO3_MOUSE |
| 24 | 27 | 87.1 | 522 | 1 | STA_RICCO |
| 25 | 27 | 87.1 | 555 | 1 | PERF HUMAN |
| 26 | 27 | 87.1 | 802 | 1 | CSD2_ECOLI |
| 27 | 27 | 87.1 | 839 | 1 | CSD1_ECOLI |
| 28 | 27 | 87.1 | 1025 | 1 | PMPF_CHLMU |
| 29 | 27 | 87.1 | 1034 | 1 | PMPF_CHLTR |
| 30 | 26 | 83.9 | 116 | 1 | HV45_MOUSE |
| 31 | 26 | 83.9 | 149 | 1 | NDK_ANASP |
| 32 | 26 | 83.9 | 151 | 1 | NDK_SYNEX |
| 33 | 26 | 83.9 | 196 | 1 | YB05_THEAC |

| | | | | | |
|----|----|------|-----|---|------------|
| 34 | 26 | 83.9 | 240 | 1 | DAPB_STAAM |
| 35 | 26 | 83.9 | 240 | 1 | DAPB_STAAM |
| 36 | 26 | 83.9 | 240 | 1 | DAPB_STAAM |
| 37 | 26 | 83.9 | 240 | 1 | DAPB_STAAM |
| 38 | 26 | 83.9 | 250 | 1 | TRYP_PLEPL |
| 39 | 26 | 83.9 | 301 | 1 | OMPG_ECOLI |
| 40 | 26 | 83.9 | 382 | 1 | DICH_DROME |
| 41 | 26 | 83.9 | 390 | 1 | P53_MOUSE |
| 42 | 26 | 83.9 | 391 | 1 | P53_RAT |
| 43 | 26 | 83.9 | 395 | 1 | YUGN_SALTY |
| 44 | 26 | 83.9 | 396 | 1 | PPAS_XENLA |
| 45 | 26 | 83.9 | 408 | 1 | O32A_DROME |

ALIGNMENTS

RESULT 1

| | | | | |
|----|--|-----------|------|---------|
| ID | CU08_LOCM1 | STANDARD; | PRT; | 148 AA. |
| AC | P11734; | | | |
| DT | 01-OCT-1989 (Rel. 12, Created) | | | |
| DT | 01-OCT-1989 (Rel. 12, Last sequence update) | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | |
| DE | Cuticle protein 8 (IM-8) (IM-ACP 8) | | | |
| OS | Locusta migratoria (Migratory locust) | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea; | | | |
| OC | Acridoidea; Acrididae; Oedipodinae; Locusta. | | | |
| OX | NCBI_TaxID=7004; | | | |
| RN | [1] | | | |
| RP | SEQUENCE | | | |
| RX | MEDLINE=90073593; PubMed=2590176; | | | |
| RA | Klarskov K., Hoejrup P., Andersen S.O., Roepstorff P.; | | | |
| RT | "Plasma-desorption mass spectrometry as an aid in protein sequence determination. Application of the method on a cuticular protein from the migratory locust (Locusta migratoria)."; | | | |
| RL | Biochem. J. 262:923-930(1989). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 1-56. | | | |
| RX | MEDLINE=86108304; PubMed=3943519; | | | |
| RA | Hoejrup P., Andersen S.O., Roepstorff P.; | | | |
| RT | "Isolation, characterization, and N-terminal sequence studies of cuticular proteins from the migratory locust, Locusta migratoria."; | | | |
| RL | Eur. J. Biochem. 154:153-159(1986). | | | |
| CC | !- FUNCTION: Component of the cuticle of migratory locust which contains more than 100 different structural proteins. | | | |
| CC | !- DOMAIN: The tetrapeptide (A-A-P-[AV]) repeats found throughout the protein are also present in many proteins constituting the protective envelope of other species. | | | |
| CC | !- SIMILARITY: Contains 1 cuticle consensus domain. | | | |
| DR | PIR; S05638; S05638. | | | |
| DR | InterPro; IPR000618; Insect cuticle. | | | |
| DR | Pfam; PF00379; Chitin_bind_4; 1. | | | |
| DR | PRINTS; PR00947; CUTICLE. | | | |
| DR | PROSITE; PS00233; CUTICLE; 1. | | | |
| KW | Structural protein; Cuticle; Repeat. | | | |
| FT | REPEAT 16 19 | | | |
| FT | REPEAT 22 22 | | | |
| FT | REPEAT 28 31 | | | |
| FT | REPEAT 37 40 | | | |
| FT | REPEAT 44 47 | | | |
| SQ | SEQUENCE 148 AA; 15224 MW; C6EADC27C593ACE6 CRC64; | | | |

Query Match 100.0%; Score 31; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYGVH 5
Db 64 NYGVH 68

```

RESULT 2
PT56_YEAST
ID PT56_YEAST STANDARD; PRT; 412 AA.
AC P25270;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ribosome methyltransferase PT56 (EC 2.1.1.-) (Mitochondrial large
DE ribosomal RNA ribose methylase) (21S rRNA [GM2251] 2'-O-
DE methyltransferase).
GN PT56 OR YOR201C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94090319; PubMed=8266080;
RA Sirum-Connolly K, Mason T.L.;
RT "Functional requirement of a site-specific ribose methylation in
RT ribosomal RNA.";
RL Science 262:1886-1889(1993).
RN [2]
RN SEQUENCE FROM N.A.
RA Hughes B., Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 1-95 FROM N.A.
RP RX MEDLINE=86093663; PubMed=3001645;
RA Struhl K.;
RT "Nucleotide sequence and transcriptional mapping of the yeast
RT pet56-his3-dedi gene region.";
RL Nucleic Acids Res. 13:887-860(1985).
CC -!- FUNCTION: Encodes an essential ribose methyltransferase that
CC specifically modifies to 2'-O-methylguanosine a universally
CC conserved nucleotide, G-2270, in the peptidyl transferase center
CC of the mitochondrial large ribosomal RNA (21S). This modification
CC seems to be important for the normal accumulation of this latter
CC 21S subunit.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: TO S.POMBE SPBC1347.13C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L19947; AAA74564.1; -
CC EMBL; X03245; CAA27002.1; -
CC EMBL; 275107; CAA39414.1; -
CC PIR; S48881; S48881.
CC Germonline; 143789; -.
CC SGD; S0005727; PET56.
CC GO; GO:0005739; C:mitochondrion; IDA.
CC GO; GO:0008989; F:rRNA (guanine-N1-)-methyltransferase activity; IDA.
CC GO; GO:000154; P:rRNA modification; IDA.
CC InterPro; IPR004441; rRNA methyl 3.
CC InterPro; IPR001537; Spou_methylase.
CC Pfam; PF00588; Spou_methylase; 1.
CC ProDom; PD001243; Spou_methylase; 1.
CC TIGRPFAMs; TIGR00196; rRNA_methyl 3; 1.
CC Mitochondrion; Transferase; Methyltransferase.
KW Mitochondrion; Transferase; Methyltransferase.
SQ SEQUENCE 412 AA; 46387 MW; C01B10254C0E0EA8 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
|||||
Db 215 NYGVH 219

RESULT 4
MYM2 HUMAN
ID MYM2 HUMAN STANDARD; PRT; 1465 AA.
AC P54236;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myomesin 2 (M-protein) (165 kDa titin-associated protein) (165 kDa
DE connectin-associated protein).
GN MYOM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP RX TISSUE=Skeletal muscle;
RX MEDLINE=94095665; PubMed=7505783;
RA Vinkemeier U., Obermann W., Weber K., Fuerst D.O.;

Query Match 100.0%; Score 31; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
|||||
Db 215 NYGVH 219

RESULT 3
YYS3 CAEEL
ID YYS3 CAEEL STANDARD; PRT; 435 AA.
AC P49191;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 51.5 kDa protein D2024.3 in chromosome IV.
GN D2024.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Du Z., Gattung S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the ELO family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U41011; AAA82288.1; -
CC PIR; T34200; T34200.
CC WormPep; D2024.3; CE04292.
CC InterPro; IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS01188; ELO; 1.
CC Hypothetical protein; Transmembrane.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 239 259 POTENTIAL.
FT TRANSMEM 278 298 POTENTIAL.
SQ SEQUENCE 435 AA; 51480 MW; F7940A74A1969914 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
|||||
Db 215 NYGVH 219
```

"The globular head domain of titin extends into the center of the sarcomeric M band. cDNA cloning, epitope mapping and immunoelectron microscopy of two titin-associated proteins.",
 J. Cell Sci. 106:319-330(1993).
 CC -!- FUNCTION: Major component of the vertebrate myofibrillar M band.
 CC Binds myosin, titin, and light meromyosin. This binding is dose dependent.
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -!- SIMILARITY: Contains 5 fibronectin type III domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X69089; CAA48832.1; ..
 CC PIR; S43529; S43529.
 CC HSSP; P40189; IBOU.
 CC Genew; HGNC:7614; MYOM2.
 CC MIM; 603509; ..
 CC GO; GO:0008307; P: structural constituent of muscle; TAS.
 CC GO; GO:0006936; P: muscle contraction; TAS.
 CC InterPro; IPR008957; FN III-like.
 CC InterPro; IPR003961; FN III.
 CC InterPro; IPR003962; FNIII subd.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_C2.
 CC Pfam; PF00041; In3; 5.
 CC Pfam; PF00047; Ig; 4.
 CC PRINTS; PR00014; FNTYPEIII.
 CC SMART; SM0060; FN3; 3.
 CC SMART; SM00408; IGC2; 1.
 CC PROSITE; PS50835; IGC2; 1.
 KW Immunoglobulin domain; Muscle protein; Thick filament; Repeat.
 FT DOMAIN 154 245
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT FIBRONECTIN TYPE-III 1.
 FT FIBRONECTIN TYPE-III 2.
 FT FIBRONECTIN TYPE-III 3.
 FT FIBRONECTIN TYPE-III 4.
 FT FIBRONECTIN TYPE-III 5.
 FT FIBRONECTIN TYPE-III 5.
 FT IG-LIKE C2-TYPE 3.
 FT IG-LIKE C2-TYPE 4.
 FT IG-LIKE C2-TYPE 5.
 FT IG-LIKE C2-TYPE 5.
 SQ SEQUENCE 1465 AA; 164793 MW; 6F2927B5A1E69F2D CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 1465;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 DB 208 NYGVH 212
 RESULT 5
 TYCC BREPA STANDARD; PRT; 6486 AA.
 AC Q30409;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrocidine synthetase III [includes: ATP-dependent asparagine
 DE adenylation synthetase III (AsnA) (Asparagine synthetase); ATP-dependent glutamine
 DE adenylation synthetase III (GlnA) (Glutamine synthetase); ATP-dependent tyrosine
 DE adenylation synthetase III (TyrA) (Tyrosine synthetase); ATP-dependent valine
 DE (ValA) (Valine synthetase); ATP-dependent ornithine adenylation synthetase
 DE (OrnA) (Ornithine synthetase); ATP-dependent leucine adenylation synthetase (LeuA)
 DE (LeuA) (Leucine synthetase)].
 GN TYCC.

OS Brevibacillus parabrevis.
 CC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
 CC NCBI_TaxID=54914;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 8185 / IAM 1031 / IFO 3331 / NCD0 717 / NCIB 8598;
 RX MEDLINE=98012987; PubMed=9352938;
 RA Meotz H.D.; Marahiel M.A.;
 RT "The tyrocidine biosynthesis operon of Bacillus brevis: complete
 RT nucleotide sequence and biochemical characterization of functional
 RT internal adenylation domains.";
 RL J. Bacteriol. 179:6843-6850(1997).
 CC -!- FUNCTION: INCORPORATES SIX AMINO ACIDS (FOR TYROCIDINE A, ASN,
 CC GLN, TYR, VAL, ORN, AND LEU) IN THEIR L-CONFIGURATION INTO THE
 CC PEPTIDE PRODUCT.
 CC -!- COFACTOR: Contains 6 covalently bound phosphopantetheines (By
 CC similarity).
 CC -!- PATHWAY: Cyclic peptide antibiotic tyrocidine biosynthesis.
 CC -!- SUBUNIT: LARGE MULTIMERIC COMPLEX OF TYCA, TYCB AND TYCC.
 CC -!- DOMAIN: CONSISTS OF SIX MODULES, AND HARBORS A PUTATIVE
 CC THIOESTERASE DOMAIN AT ITS C-TERMINAL END. EACH MODULE
 CC INCORPORATES ONE AMINO ACID INTO THE PEPTIDE PRODUCT AND CAN BE
 CC FURTHER SUBDIVIDED INTO DOMAINS RESPONSIBLE FOR SUBSTRATE
 CC ADENYLATION, THIOLEAD, CONDENSATION (NOT FOR THE INITIATION
 CC MODULE), AND EPIMERIZATION (OPTIONAL), AND N METHYLATION
 CC (OPTIONAL).
 CC -!- MISCELLANEOUS: TYROCIDINE IS A MIXTURE OF FOUR CYCLIC
 CC DECAPEPTIDES, TYROCIDINE A (D-PHE-PRO-PHE-D-PHE-ASN-GLN-TYR-VAL-
 CC ORN-LEU), B, C, AND D, IN WHICH PHE, AT POSITIONS 3, 4, AND TYR
 CC RESIDUES ARE GRADUALLY REPLACED BY TRP, DEPENDING ON THE RELATIVE
 CC CONCENTRATIONS OF THESE AMINO ACIDS IN THE GROWTH MEDIUM.
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC -!- SIMILARITY: Contains 6 acyl carrier domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF004835; AAC45930.1; ..
 CC PDB; 1DNY; 17-MAY-00.
 CC InterPro; IPR000873; AMP-bind.
 CC InterPro; IPR001242; Condensatn.
 CC InterPro; IPR006163; Pp_bind.
 CC InterPro; IPR006162; Ppantne.S.
 CC InterPro; IPR000379; Ser_estrs.
 CC InterPro; IPR001031; Thioesterase.
 CC Pfam; PF00501; AMP-binding; 6.
 CC Pfam; PF00668; Condensation; 6.
 CC Pfam; PF00550; Pp-binding; 6.
 CC Pfam; PF00975; Thioesterase; 1.
 CC PRINTS; PR00154; AMPBINDING.
 CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 6.
 CC PROSITE; PS00455; AMP BINDING; 6.
 CC PROSITE; PS00075; ACP DOMAIN; 6.
 KW Ligase; Antibiotic biosynthesis; Phosphopantetheine;
 KW Multifunctional enzyme; Repeat; 3D-structure.
 FT REPEAT 466 1038
 FT DOMAIN 1 (ASPARAGINE-ACTIVATING).
 FT REPEAT 1521 2070
 FT DOMAIN 2 (GLUTAMINE-ACTIVATING).
 FT REPEAT 2536 3113
 FT DOMAIN 3 (TYROSINE-ACTIVATING).
 FT REPEAT 3590 4149
 FT DOMAIN 4 (VALINE-ACTIVATING).
 FT REPEAT 4606 5203
 FT DOMAIN 5 (ORNITHINE-ACTIVATING).
 FT REPEAT 5658 6245
 FT DOMAIN 6 (LEUCINE-ACTIVATING).
 FT DOMAIN 970 1037
 FT ACYL CARRIER (ACP) 1.
 FT DOMAIN 2007 2074
 FT ACYL CARRIER (ACP) 2.
 FT DOMAIN 3045 3112
 FT ACYL CARRIER (ACP) 3.
 FT DOMAIN 4080 4147
 FT ACYL CARRIER (ACP) 4.
 FT DOMAIN 5124 5191
 FT ACYL CARRIER (ACP) 5.
 FT DOMAIN 6167 6234
 FT ACYL CARRIER (ACP) 6.

FT BINDING 1000 1000 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 2037 2037 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 3075 3075 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 4110 4110 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 5154 5154 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 6197 6197 PHOSPHOPANTETHEINE (BY SIMILARITY).
 SQ SEQUENCE 6486 AA; 724011 MW; 4934900A07DF786 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 6486;

Best Local Similarity 100.0%; Pred. No. 2.2e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0;

Qy 1 NYGVH 5

Db 2664 NYGVH 2668

RESULT 6

ID CARA LEPIN STANDARD; PRT; 363 AA.

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase glutamine chain).
 GN CARA OR LA1239.

OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.

OX NCBI_TaxID=173;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=56601 / Serogroup Interchaemorrhagiae / Seroovar lai;

RX MEDLINE=22598143; PubMed=12712204;

RA REN S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,

RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,

RA Jiang Y.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,

RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,

RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,

RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,

RA Xu J.-G., Zhao G.-P.

RT "Unique physiological and pathogenic features of Leptospira

interrogans revealed by whole-genome sequencing.";

RL Nature 422:888-893(2003).

CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +

phosphate + glutamate + carbamoyl phosphate.

CC -!- PATHWAY: Arginine biosynthesis.

CC -!- SUBUNIT: Pyrimidine biosynthesis; first step.

CC promotes the hydrolysis of glutamine to ammonia which is used by

the large (or ammonia) chain to synthesize carbamoyl phosphate (By

similarity).

CC -!- SIMILARITY: Belongs to the cara family.

CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE011305; AAN48438.1; --

DR HANAP; MF_01209; --; 1.

DR InterPro; IPR006274; CarA_synth_small.

DR InterPro; IPR001317; CP_synthGATase.

DR InterPro; IPR002474; CP_synthsmall.

DR InterPro; IPR000991; GATase 1.

DR Pfam; PF00988; CPsase_sm_chain; 1.

DR Pfam; PF00117; GATase_1.

DR PRINTS; PR00099; CPsGATase.

DR PRINTS; PR00096; GATase.

DR TIGRFAMS; TIGR01368; CPsaseIsmall; 1.
 DR PROSITE; PS00442; GATASE TYPE I; 1.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
 KW Glutamine amidotransferase; Complete proteome.
 FT DOMAIN 1 173 CFSASE.
 FT DOMAIN 174 363 GLUTAMINE AMIDOTRANSFERASE.
 FT ACT_SITE 253 253 GATASE (BY SIMILARITY).
 SQ SEQUENCE 363 AA; 40261 MW; 58BD2BB8F43D0BEE CRC64;

Query Match 96.8%; Score 30; DB 1; Length 363;

Best Local Similarity 80.0%; Pred. No. 22;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5

Db 61 NYGIH 65

RESULT 7

ID PUCJ_BACSU STANDARD; PRT; 449 AA.

AC O32139;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Uric acid permease pucd.

GN PUCJ OR BSU32430.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Beresford R., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borries V., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruch C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Grandi G.,

RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porvollik S., Prescott A.M.,

RA Pressac E., Fujic P., Furnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Takeuchi M., Takashi H., Takashi H., Takemaru K.,

RA Takeuchi M., Takatoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Terasaki V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H., Zumstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

subtilis";

RL Nature 390:249-256(1997).

RN [2]

RP FUNCTION.

RC STRAIN=168;

RX MEDLINE=21242727; PubMed=11344136;

RA Schultz A.C., Nygaard P., Saxild H.H.;

RT "Functional analysis of 14 genes that constitute the purine catabolic

pathway in Bacillus subtilis and evidence for a novel regulon

controlled by the PucR transcription activator.";

RL J. Bacteriol. 183:3293-3302(2001).
 CC -1- FUNCTION: Uptake of uric acid.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- INDUCTION: Expression is very low in excess nitrogen (glutamate
 CC plus ammonia). and is induced during limiting-nitrogen conditions
 CC (glutamate). Expression is further induced when allantoin or uric
 CC acid are added during limiting-nitrogen conditions.
 CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z99120; CAB15233.1; --
 CC PIR; E70016; E70016.
 CC Subtilist; BG13984; pucJ.
 CC InterPro; IPR006042; Xan ur permease.
 CC InterPro; IPR006043; XanT/urac/vitC.
 CC Pfam; PF00860; xan ur permease; 1.
 CC TIGRfam; TIGR00801; ncs2; 1.
 CC PROSITE; PS01116; XANTH URACIL PERMEASE; 1.
 CC Transmembrane; Transport; Complete proteome.
 CC TRANSMEM 11 31 POTENTIAL.
 CC TRANSMEM 41 61 POTENTIAL.
 CC TRANSMEM 67 87 POTENTIAL.
 CC TRANSMEM 91 111 POTENTIAL.
 CC TRANSMEM 119 139 POTENTIAL.
 CC TRANSMEM 158 178 POTENTIAL.
 CC TRANSMEM 191 211 POTENTIAL.
 CC TRANSMEM 229 249 POTENTIAL.
 CC TRANSMEM 277 297 POTENTIAL.
 CC TRANSMEM 313 333 POTENTIAL.
 CC TRANSMEM 334 354 POTENTIAL.
 CC TRANSMEM 372 392 POTENTIAL.
 CC TRANSMEM 401 421 POTENTIAL.
 CC SEQUENCE 449 AA; 47100 MW; 9B97CC42330C087 CRC64;

Query Match 96.8%; Score 30; DB 1; Length 449;
 Best Local Similarity 80.0%; Pred. No. 27;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||:|
 Db 90 NYGIH 94

RESULT 8

AMYL1 SCHPO STANDARD; PRT; 581 AA.
 AC Q09840;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable alpha-amylase C23D3.14c precursor (EC 3.2.1.1) (1,4-alpha-D-
 DE glucan glucanohydrolase).
 GN SPAC23D3.14c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetes; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=2184801; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown J., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Dominguez A., Revuelta J.L., Sanchez M., del Rey F., Benito J.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe."
 RT Nature 415:871-880(2002).
 RL -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: Binds 2 calcium ions per subunit. Calcium is inhibitory
 CC at high concentrations (By similarity).
 CC -1- SIMILARITY: Belongs to family 13 of Glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z64354; CAA91249.1; --
 CC PIR; T38289; S62505.
 CC HSP; P10529; 7TAA.
 CC GeneDB Spombe; SPAC23D3.14c; --
 CC InterPro; IPR00589; Alpha_ami1_cat_sub.
 CC InterPro; IPR006047; Alpha_ami1_cat.
 CC Pfam; PF00128; alpha-amylase; 1.
 CC SMART; SMO0642; Amy; 1.
 CC KW Hypothetical protein; Signal; Carbohydrate metabolism; Hydrolase; Glycosidase;
 KW Calcium-binding; Signal; Glycoprotein.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 581 PROBABLE ALPHA-AMYLASE C23D3.14C.
 FT ACT_SITE 233 233 BY SIMILARITY.
 FT ACT_SITE 237 237 BY SIMILARITY.
 FT ACT_SITE 325 325 BY SIMILARITY.
 FT METAL 147 147 CALCIUM 1 (BY SIMILARITY).
 FT METAL 202 202 CALCIUM 1 (BY SIMILARITY).
 FT METAL 233 233 CALCIUM 2 (BY SIMILARITY).
 FT METAL 237 237 CALCIUM 1 (VIA CARBONYL OXYGEN) (BY
 FT SIMILARITY).
 FT METAL 257 257 CALCIUM 2 (BY SIMILARITY).
 FT DISULFID 56 64 BY SIMILARITY.
 FT DISULFID 176 191 BY SIMILARITY.
 FT DISULFID 267 311 BY SIMILARITY.
 FT CARBOHYD 291 291 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 581 AA; 67004 MW; FE9DE99D323E1890 CRC64;

Query Match 96.8%; Score 30; DB 1; Length 581;
 Best Local Similarity 80.0%; Pred. No. 35;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 435 NYGIH 439

RESULT 9
KALM_HUMAN
ID KALM_HUMAN STANDARD; PRT; 680 AA.
AC P23352;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Anosmin 1 precursor (Kallmann syndrome protein) (Adhesion molecule-like X-linked).
DE like X-linked).
GN KAL1 OR KAL OR ADMLX OR KALIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92005720; PubMed=1913827;
RA Legouis R., Hardelin J.-P., Levilliers J., Claverie J.-M., Compain S., Wunderle V., Millasseau P., le Paslier D., Cohen D., Caterina D., Bouquelieret L., Delenarre-Van de Waal H., Lutfalla G., Weissenbach J., Petit C.;
RA "The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules.";
RT Cell 67:423-435 (1991).
RL [2]
RN REVISIONS.
RP MEDLINE=93265164; PubMed=1303284;
RX del Castillo I., Cohen-Salmon M., Blanchard S., Lutfalla G., Petit C.;
RA "Structure of the X-linked Kallmann syndrome gene and its homologous pseudogene on the Y chromosome.";
RT Nat. Genet. 2:305-310 (1992).
RL [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=92018217; PubMed=1922361;
RA Franco B., Guoli S., Pragliola A., Inceri B., Bardoni B., Tonlorenzi R., Carrozo R., Maestrini E., Pieretti M., Lawrence C., Tailon-Willer P., Brown C.J., Willard H.F., Ballabio A., Persico N.G., Camerino G., Ballabio A.;
RA "A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules.";
RT Nature 353:529-536 (1991).
RL [4]
RN SEQUENCE OF 1-71 FROM N.A.
RX MEDLINE=96069588; PubMed=7590336;
RA Cohen-Salmon M., Tronche F., del Castillo I., Petit C.;
RA "Characterization of the promoter of the human KAL gene, responsible for the X-chromosome-linked Kallmann syndrome.";
RT Gene 164:235-242 (1995).
RL [5]
RN CHARACTERIZATION.
RX MEDLINE=96429296; PubMed=8932397;
RA Soussi-Yanicostas N., Hardelin J.-P., del Mar Arroyo-Jimenez M., Arduin O., Legouis R., Levilliers J., Traincard F., Betton J.-M., Cabanie L., Petit C.;
RA "Initial characterization of anosmin-1, a putative extracellular matrix protein synthesized by definite neuronal cell populations in the central nervous system.";
RT J. Cell Sci. 109:1749-1757 (1996).
RL [6]
RN VARIANT KAL1 LYS-267, AND VARIANT VAL-534.
RX MEDLINE=93278384; PubMed=8504298;
RA Hardelin J.-P., Levilliers J., Blanchard S., Carel J.-C., Leuenegger M., Pinard-Berthelot J.-P., Bouleux P., Petit C.;
RA "Heterogeneity in the mutations responsible for X chromosome-linked Kallmann syndrome.";
RT Hum. Mol. Genet. 2:373-377 (1993).
RL [7]
RN VARIANT KAL1 LYS-514, AND VARIANT VAL-534.
RX MEDLINE=98251583; PubMed=9589672;
RA Maya-Nunez G., Zenteno J.C., Ulloa-Aguirre A., Kofman-Alfaro S., Mendez J.P.;
RA "A recurrent missense mutation in the KAL gene in patients with

X-linked Kallmann's syndrome.";
J. Clin. Endocrinol. Metab. 83:1650-1653 (1998).
-i- FUNCTION: May be an adhesion-like molecule with anti-protease activity.
-i- SUBCELLULAR LOCATION: Secreted. Localized at cell surface.
-i- PTM: N-glycosylated.
-i- DISEASE: Defects in KAL1 is the cause of Kallmann syndrome 1 (KAL1) [MIM:308700]; also abbreviated KS. KAL1 is a genetic disorder that associates hypogonadotropic hypogonadism and anosmia. In this disease, the normal embryonic migration of GNRH-synthesizing neurons from the olfactory placodes to the hypothalamic region as well as the axonal extension of olfactory neurons towards the forebrain are impaired.
-i- SIMILARITY: Contains 4 fibronectin type III domains.
-i- SIMILARITY: Contains 1 WAP-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M97252; AAA59202.1; -
DR ENBL; S60085; AAB20108.1; ALT_SEQ.
EMBL; X60299; CAA42841.1; -
EMBL; X82034; CAA57554.1; -
PIR; S17982; S17982.
Genew; HGNC:6211; KAL1.
MIM; 308700; -
GO; GO:0005578; C:extracellular matrix; TAS.
GO; GO:0005615; C:extracellular space; TAS.
GO; GO:0005201; F:extracellular matrix structural constituent; TAS.
GO; GO:0007411; P:axon guidance; TAS.
GO; GO:0007155; P:cell adhesion; TAS.
GO; GO:0006928; P:cell motility; TAS.
GO; GO:0006935; P:chemotaxis; TAS.
InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003951; FN_III.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00095; wap; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4 DISULFIDE CORE; 1.
KW Cell adhesion; Glycoprotein; Serine protease inhibitor; Repeat; Signal; Polymorphism; Disease mutation.
FT SIGNAL 1 20
FT CHAIN 21 680 ANOSMIN 1.
FT DOMAIN 21 120 "CYSTEINE BOX".
FT DOMAIN 130 176 FIBRONECTIN TYPE-III 1.
FT DOMAIN 181 285 FIBRONECTIN TYPE-III 2.
FT DOMAIN 286 402 FIBRONECTIN TYPE-III 3.
FT DOMAIN 403 540 FIBRONECTIN TYPE-III 4.
FT DOMAIN 541 661 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC... (POTENTIAL).
FT VARIANT 267 267 N -> K (in KAL1).
FT FTID=VAR_007720.
FT VARIANT 514 514 E -> K (in KAL1).
FT FTID=VAR_012742.
FT VARIANT 534 534 I -> V.
FT FTID=VAR_007721.
FT NN -> VR (in REF. 4).
FT CONFLICT 70 71
FT CONFLICT 373 373 E -> K (in REF. 3).
FT CONFLICT 540 540 A -> R (in REF. 3).
FT SEQUENCE 680 AA; 76065 MW; 5D6ACC9F14B5F5F8 CRC64;
SQ

Query Match 96.8%; Score 30; DB 1; Length 680;
 Best Local Similarity 80.0%; Pred. No. 41;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 DB 224 NYGIH 228

RESULT 10
 PREL YEAST STANDARD; PRT; 686 AA.
 AC P22791;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 10-OCT-1993 (Rel. 27, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ferric reductase transmembrane component 1 precursor (BC 1.16.1.7)
 DE Ferric-chelate reductase 1.
 GN PRE1 OR YLR1214W OR L8167.2
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F113.
 RX MEDLINE=92237270; PubMed=1570306;
 RA Dancis A., Roman D.G., Anderson G.J., Hinnebusch A.G., Klausner R.D.;
 RA "Ferric reductase of Saccharomyces cerevisiae: molecular
 RA characterization, role in iron uptake, and transcriptional control by
 RA iron.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:3869-3873 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
 RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.-J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
 RA Mueller-Auer S., Newnich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetle D., Fumelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urescarazu L.A., Vandenbol M., Verhasselt P.,
 RA Vierendeels F., Voet M., Volckert G., Voss H., Wambutt R., Kedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Fani J., Hohnes J.D.;
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:87-90 (1997).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93057491; PubMed=1431884;
 RA Anderson G.J., Lesuisse E., Dancis A., Roman D.G., Labbe P.,
 RA Klausner R.D.;
 RA "Ferric iron reduction and iron assimilation in Saccharomyces
 RA cerevisiae.";
 RL J. Inorg. Biochem. 47:249-255 (1992).
 CC -!- FUNCTION: Reductase activity that acts on ferric iron chelates
 CC external to the cell. Plays a role in iron uptake. May
 CC participate in the transport of electrons from cytoplasm to an
 CC extracellular substrate (ferric ion) via FAD and heme
 CC intermediates. May also participate in Cu(II) reduction and Cu(I)
 CC uptake.
 CC -!- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
 CC -!- COFACTOR: FAD (Probable).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- INDUCTION: BY IRON DEPRIVATION. REPRESSED BY IRON UPTAKE.
 CC -!- SIMILARITY: Belongs to the FRE / CYBB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; M86908; AAA34608.1; -;
 CC EMBL; U14913; AAB67424.1; -;
 CC PIR; S30075; S30075.
 CC Germonline; 142276; -;
 CC SGD; S0004204; FRE1.
 CC GO; GO:0005886; C:plasma membrane; IDA.
 CC GO; GO:0000293; F:ferric-chelate reductase activity; IDA.
 CC GO; GO:0035677; P:copper ion import; IDA.
 CC GO; GO:0006826; P:iron ion transport; IDA.
 CC Pfam; PF01794; Ferric_reduct; 1.
 CC InterPro; IPR002916; Ferric_reduct.
 CC Oxidoreductase; Electron transport; Transmembrane; Iron transport;
 KW FAD; NAD; Copper; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 22 POTENTIAL
 FT CHAIN 23 686 FERRIC REDUCTASE TRANSMEMBRANE COMPONENT
 FT 1.
 FT NP_BIND 462 468 FAD (POTENTIAL).
 FT NP_BIND 532 540 NAD (POTENTIAL).
 FT TRANSMEM 147 169 1 (POTENTIAL).
 FT TRANSMEM 216 236 2 (POTENTIAL).
 FT TRANSMEM 258 277 3 (POTENTIAL).
 FT TRANSMEM 298 316 4 (POTENTIAL).
 FT TRANSMEM 329 348 5 (POTENTIAL).
 FT TRANSMEM 369 397 6 (POTENTIAL).
 FT TRANSMEM 529 550 7 (POTENTIAL).
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 686 AA; 78853 MW; 7F6BB3B93A95D6A3 CRC64;

Query Match 96.8%; Score 30; DB 1; Length 686;
 Best Local Similarity 80.0%; Pred. No. 41;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 DB 135 NYGIH 139

RESULT 11
 HV3J HUMAN STANDARD; PRT; 121 AA.
 AC P01771;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V-III region HIL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79124695; PubMed=420800;
 RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
 RA "Amino acid sequence of the VH region of human myeloma
 RA cryoimmunoglobulin IGG Hil.";
 RL Biochemistry 18:553-560 (1979).
 CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 CC PROTEIN.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A02054; GLHUHL.
 DR HSSP; P01772; 2PB4.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding; NAS.

DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003596; IG_v.
 DR Pfam: PF00047; ig_1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KW Immunoglobulin v region; Pyrrolidone carboxylic acid.
 FT DOMAIN 1 112
 FT MOD_RES 1 1
 FT NON_TER 121 121
 SQ SEQUENCE 121 AA; 13566 MW; 480FC53610EFD5DAB CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 121;
 Best Local Similarity 80.0%; Pred. No. 21;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 31 NYGMH 35
 RESULT 12
 ID NUIM CAEEL STANDARD; PRT; 212 AA.
 AC Q22619;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NADH-ubiquinone oxidoreductase 23 kDa subunit, mitochondrial precursor
 DE (EC 1.6.5.3) (EC 1.6.9.3) (Complex I-23KD) (CI-23KD).
 GN T20H4.5
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=93377169; PubMed=10446229;
 RA Hough R.F., Lingam A.T., Bass B.L.;
 RT "Caenorhabditis elegans mRNAs that encode a protein similar to ADARs
 RT derive from an operon containing six genes."
 RL Nucleic Acids Res. 27:3424-3432(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 Du Z.;
 RA Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RL
 CC -!- FUNCTION: May donate electrons to ubiquinone.
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -!- COFACTOR: Binds 2 4Fe-4S clusters per subunit (By similarity).
 CC -!- SUBUNIT: Complex I is composed of 45 different subunits (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the complex I 23 kDa subunit family.
 CC -!- SIMILARITY: The iron-sulfur centers are similar to those of
 CC bacterial-type 4Fe-4S ferredoxins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF140272; AAD34863.1;
 CC EMBL; U00037; AA050662.1;
 CC PIR; T16914; T16914.
 CC HSSP; P00198; 2FDN.
 CC WormPep; T20H4.5; CE00832.
 DR InterPro: IPR001450; 4Fe4S-ferredoxin.
 DR Pfam; PF00037; fer4; 2.
 DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transit peptide;
 KW Iron-sulfur; 4Fe-4S.
 FT TRANSIT 1 ? MITOCHONDRION (BY SIMILARITY).
 FT CHAIN ? 212 NADH-UBIQUINONE OXIDOREDUCTASE 23 KDA
 FT SUBUNIT.
 FT METAL 113 113 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 116 116 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 119 119 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 123 123 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 152 152 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 155 155 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 158 158 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 162 162 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 212 AA; 23870 MW; F50C413D89E80AB7 CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 212;
 Best Local Similarity 80.0%; Pred. No. 37;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 55 NYGLH 59
 RESULT 13
 ID NDF6 HUMAN STANDARD; PRT; 337 AA.
 AC Q96NR8; Q9H3H6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 6 (NeuroD6) (My051 protein).
 DE NEUROD6.
 GN NEUROD6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=fetal brain;
 RA Mac Y.M., Xie Y., Zheng Z.H.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie K.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatasuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
 RA Masuno Y., Nagai K., Isogai T.;
 RL "NEDO human cDNA sequencing project."
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
 CC with E47. May be a trans-acting factor involved in the development
 CC and maintenance of the mammalian nervous system. Transactivates
 CC the promoter of its own gene (By similarity).
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
 CC in positions 300 and 307.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

or send an email to license@isb-sib.ch .
-----
CC EMBL; AF063609; AAG43167.1; ALT_FRAME.
CC EMBL; AK055238; BAB70885.1; -.
CC DR EMBL; AK055238; BAB70885.1; -.
CC DR Genew; HGNC:13804; NEUROD6.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS50888; HLH; 1.
CC KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT FT DOMAIN 54 63 POLY-GLU.
FT FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT FT DNA BIND 95 106 BASIC DOMAIN
FT FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 337 AA; 38705 MW; 680F4127AC0F809E CRC64;
-----
Query Match 90.3%; Score 28; DB 1; Length 337;
Best Local Similarity 80.0%; Pred. No. 58;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 NYGVH 5
Db 286 NYGMH 290
-----
RESULT 14
IDF6 MOUSE
IDF6_MOUSE STANDARD; PRT; 337 AA.
AC P48986;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 6 (Neuro6) (Atonal protein homolog
DE 2) (Helix-loop-helix protein MATH-2) (MATH2) (NEX-1 protein).
DE NEUROD6 OR ATOH2 OR ATH2 OR NEX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RC MEDLINE=95262673; PubMed=7744035;
RA Shimizu C., Akazawa C., Nakanishi S., Kageyama R.;
RT "MATH-2", a mammalian helix-loop-helix factor structurally related
RT to the product of Drosophila proneural gene atonal, is specifically
RT expressed in the nervous system";
RL Eur. J. Biochem. 229:239-248(1995).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RC MEDLINE=95200803; PubMed=7545978;
RA Bartholomae A., Nave K.-A.;
RT "NEX-1: a novel brain-specific helix-loop-helix protein with
RT autoregulation and sustained expression in mature cortical neurons.";
RL Mech. Dev. 48:217-228(1994).
CC -1- FUNCTION: Activates E box-dependent transcription in collaboration
CC and with E47. May be a trans-acting factor involved in the development
CC and maintenance of the mammalian nervous system. Transactivates
CC the promoter of its own gene.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Specific to the nervous system of both embryos
CC and adults. Highest levels in the cortical plate of the cerebrum.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/)

```

RT "Nucleotide sequence between the fadB gene and the rna operon from
 RT Escherichia coli."; [5]
 RL Nucleic Acids Res. 18:6439-6439(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11205551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Probst G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potanousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."; [6]
 RL Nature 409:529-533(2001).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O157:H7 / RMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Onosubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 RA Kohara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12"; [7]
 RL DNA Res. 8:11-22(2001).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157"; [8]
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -!- FUNCTION: LOW-AFFINITY POTASSIUM TRANSPORT SYSTEM. INTERACTS WITH
 CC TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA AND REQUIRES TRKE FOR
 CC TRANSPORT ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -!- SIMILARITY: Belongs to the trkH potassium transport family.
 CC -!- CAUTION: Ref.2 and Ref.4 sequences differ from that shown in
 CC position 418 onward due to frameshifts.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M87049; AAA67646.1; ALT FRAME.
 CC EMBL; AF000460; AAC76852.1; ALT FRAME.
 CC EMBL; X54687; -; NOT ANNOTATED_CDS.
 CC EMBL; AF005615; AAG59043.1; -;
 CC EMBL; AF002567; BAB36200.1; -;
 CC EMBL; AE015399; AAN45360.1; -;
 CC EMBL; AE016990; AAP18838.1; -;
 CC PIR; A91226; A91226.

DR PIR; B65190; B65190.
 DR PIR; G86072; G86072.
 DR EcoGene; EGI1021; trkH.
 DR InterPro; IPR003445; Cat transpt.
 DR InterPro; IPR004772; K transptTrk.
 DR Pfam; PF02386; TrkH; 1.
 DR TIGRfam; TIGR00933; 2a38; 1.
 DR Transmembrane; Inner membrane; Transport; Potassium transport;
 KW Complete proteome.
 FT TRANSMEM 9 29 POTENTIAL.
 FT TRANSMEM 35 55 POTENTIAL.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 109 129 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 240 260 POTENTIAL.
 FT TRANSMEM 277 297 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 397 417 POTENTIAL.
 FT TRANSMEM 423 443 POTENTIAL.
 FT TRANSMEM 423 443 MISSING (IN REF. 2).
 FT CONFLICT 144 144 V -> IV (IN REF. 4).
 FT CONFLICT 196 196 C -> S (IN REF. 4).
 FT CONFLICT 196 196 C -> S (IN REF. 4).
 SQ SEQUENCE 483 AA; 52959 MW; 1AA9CC2F83EB509A CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 483;
 Best Local Similarity 80.0%; Pred. No. 82;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 252 NYGLH 256
 Search completed: October 6, 2004, 16:30:22
 Job time : 6.07018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:59 ; Search time 5.26316 Seconds
(without alignments)
91.382 Million cell updates/sec

Title: US-09-635-974A-2

Perfect score: 31

Sequence: 1 NYGVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 31 | 100.0 | 109 | PH1026 | Ig heavy chain V r |
| 2 | 31 | 100.0 | 139 | A32456 | Ig heavy chain pre |
| 3 | 31 | 100.0 | 148 | S05638 | cuticle protein 8 |
| 4 | 31 | 100.0 | 263 | AF2424 | hypothetical prote |
| 5 | 31 | 100.0 | 325 | T33082 | hypothetical prote |
| 6 | 31 | 100.0 | 325 | B87721 | protein F56A6.1 [i |
| 7 | 31 | 100.0 | 338 | T26588 | hypothetical prote |
| 8 | 31 | 100.0 | 412 | S48981 | rRNA (guanosine-2' |
| 9 | 31 | 100.0 | 435 | T34200 | hypothetical prote |
| 10 | 31 | 100.0 | 475 | S76734 | hypothetical prote |
| 11 | 31 | 100.0 | 482 | E96500 | probable histidine |
| 12 | 31 | 100.0 | 697 | T15179 | hypothetical prote |
| 13 | 31 | 100.0 | 1465 | S43529 | 16SK protein, ske |
| 14 | 31 | 100.0 | 6486 | T34076 | tyrocidine synthet |
| 15 | 30 | 96.8 | 118 | PQ0266 | Ig heavy chain V r |
| 16 | 30 | 96.8 | 170 | AB2402 | hypothetical prote |
| 17 | 30 | 96.8 | 449 | E70016 | purine permease ho |
| 18 | 30 | 96.8 | 479 | C64655 | hypothetical prote |
| 19 | 30 | 96.8 | 479 | H71942 | probable outer mem |
| 20 | 30 | 96.8 | 581 | S63505 | probable alpha-amy |
| 21 | 30 | 96.8 | 679 | A40351 | adhesion-type prot |
| 22 | 30 | 96.8 | 680 | S17982 | Kallmann syndrome |
| 23 | 30 | 96.8 | 686 | S30075 | ferric reductase (|
| 24 | 30 | 96.8 | 788 | AF0122 | probable membrane |
| 25 | 28 | 90.3 | 121 | G1HUHL | Ig heavy chain V-1 |
| 26 | 28 | 90.3 | 140 | S70442 | Ig heavy chain pre |
| 27 | 28 | 90.3 | 196 | T28253 | hypothetical prote |
| 28 | 28 | 90.3 | 212 | T16914 | hypothetical prote |
| 29 | 28 | 90.3 | 320 | H95136 | hypothetical prote |

ALIGNMENTS

RESULT 1

PH1026

Ig heavy chain V region (clone 163-c3) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PH1026

R:Fillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH1026

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-109 <TIL>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. NO. 8.6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 31 NYGVH 35

|||||

RESULT 2

A32456

Ig heavy chain precursor V region - mouse

C:Species: Mus musculus (house mouse)

C>Date: 26-Oct-1999 #sequence_revision 31-Dec-1993 #text_change 23-Jul-1999

C:Accession: A32456

R:Dombrink-Kurtzman, M.A.; Johnson, L.S.; Riordan, G.S.; Bedzyk, W.D.; Voss Jr., E.W.

J. Biol. Chem. 264, 4513-4522, 1989

A:Title: Variable region primary structures of a high affinity anti-fluorescein immunogl

A:Reference number: A32456; MUID:89174706; PMID:2494173

A:Accession: A32456

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-139 <DM>

A:Cross-references: GB:04609; NID:G556316; PID:AAA50298.1; PID:G556317

A>Note: the authors translated the codon CAC for residue 20 as Gln, and CAC for residue

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

F:34-116/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 31; DB 2; Length 139;

Best Local Similarity 100.0%; Pred. NO. 11;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 50 NYGVH 54

RESULT 3
S05638
cuticle protein 8 - migratory locust
C:Species: Locusta migratoria (migratory locust)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: S05638; B24802
R:Klarskov, K.; Hojrup, P.; Andersen, S.O.; Roepstorff, P.
Biochem. J. 262, 923-930, 1989
A:Title: Plasma-desorption mass spectrometry as an aid in protein sequence determination
A:Reference number: S05638; MUID:90073593; PMID:2590176
A:Accession: S05638
A:Molecule type: protein
A:Residues: 1-148 <KLA>
R:Hojrup, P.; Andersen, S.O.; Roepstorff, P.
Eur. J. Biochem. 154, 153-159, 1986
A:Reference number: A91157; MUID:86108304; PMID:3943519
A:Accession: B24802
A:Molecule type: protein
A:Residues: 1-53, 'X', 55-56 <HOJ>
C:Superfamily: migratory locust cuticle protein 8
C:Keywords: structural protein
F16-19/Region: 4-residue repeat (A-A-P-[AV])
F22-25/Region: 4-residue repeat (A-A-P-[AV])
F28-31/Region: 4-residue repeat (A-A-P-[AV])
F37-40/Region: 4-residue repeat (A-A-P-[AV])
F44-47/Region: 4-residue repeat (A-A-P-[AV])

Query Match 100.0%; Score 31; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. NO. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 64 NYGVH 68

RESULT 4
AF2424
hypothetical protein all4950 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AF2424
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saeamato, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:2155285; PMID:11759840
A:Accession: AF2424
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076649.1; PID:gl7134088; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4950

Query Match 100.0%; Score 31; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. NO. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 145 NYGVH 149

RESULT 5
T33082
hypothetical protein F56A6.1 - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33082
R:Murray, J.; Rohlfing, T.; O'Neal, D.; Wilson, R.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid F56A6.
A:Reference number: Z21279
A:Accession: T33082
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-325 <MUR>
A:Cross-references: EMBL:AF067217; PIDN:AAC17014.1; GSPDB:GN00019; CESP:F56A6.1
A:Experimental source: strain Bristol N2; clone F56A6
C:Genetics:
A:Gene: CESP:F56A6.1
A:Map position: 1
A:Introns: 51/1; 135/3; 202/2; 222/2; 262/1

Query Match 100.0%; Score 31; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. NO. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 310 NYGVH 314

RESULT 6
B87721
Protein F56A6.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B87721
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites Genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B87721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <STO>
A:Cross-references: GB:chr_I; PIDN:AAC17014.1; PID:g3150502; GSPDB:GN00019; CESP:F56A6.1
A:Note: partial CDS
C:Genetics:
A:Gene: F56A6.1
A:Map position: 1

Query Match 100.0%; Score 31; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. NO. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 310 NYGVH 314

RESULT 7
T26588
hypothetical protein Y32B128.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26588
R:Gardner, A.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z20239
A:Accession: T26588
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-338 <WIL>
A:Cross-references: EMBL:AL031632; PIDN:CA921008.1; GSPDB:GN00023; CESP:Y32B128.6

A;Experimental source: clone Y32B12B
 C;Genetics:
 A;Gene: CESP:Y32B12B.6
 A;Map position: 5
 A;Introns: 28/3; 61/1; 257/1

Query Match 100.0%; Score 31; DB 2; Length 338;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 26 NYGVH 30

RESULT 8
 S48881
 rRNA (guanosine-2'-O-)-methyltransferase (EC 2.1.1.1.-) - yeast (Saccharomyces cerevisiae)
 N;Alternate names: mitochondrial large ribosomal RNA methylase; protein O4827; protein Y
 C;Species: Saccharomyces cerevisiae
 C;Date: 29-Nov-1994 #sequence_revision 12-May-1995 #text_change 21-Jul-2000
 C;Accession: S48881; S07682; S67093
 R;Sirum-Connolly, K.; Mason, T.L.
 Science 262, 1886-1889, 1993

A;Title: Functional requirement of a site-specific ribose methylation in ribosomal RNA.
 A;Accession: S48881
 A;Reference number: S48881; MUID:94090319; PMID:8266080
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-412 <STR>
 A;Cross-references: EMBL:L19947; NID:G431759; PIDN:AAA74564.1; PID:G431760
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
 R;Struhl, K.

Nucleic Acids Res. 13, 8587-8601, 1985
 A;Title: Nucleotide sequence and transcriptional mapping of the yeast pet56-his3-ded1 ge
 A;Reference number: S07681; MUID:86093663; PMID:3001645
 A;Accession: S07682
 A;Molecule type: DNA
 A;Residues: 1-95 <STR>
 A;Cross-references: EMBL:X03245; NID:G3778; PIDN:CAA27002.1; PID:G3779
 R;Hughes, B.; Pohl, T.M.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: S66685
 A;Accession: S67093
 A;Molecule type: DNA
 A;Residues: 1-412 <HUG>
 A;Cross-references: EMBL:275107; NID:G1420472; PID:e252073; PID:G1420475; MIPS:YOR201C
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: SGD:PET56
 A;Cross-references: SGD:S0005727; MIPS:YOR201C
 A;Map position: 15R
 A;Genome: nuclear
 C;Function:
 A;Description: methyltransferase; required for formation of functional mitochondrial rib
 C;Keywords: methyltransferase; mitochondrion

Query Match 100.0%; Score 31; DB 2; Length 412;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 181 NYGVH 185

RESULT 9
 T34200
 hypothetical protein D2024.3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T34200
 R;Du, Z.; Gattung, S.

submitted to the EMBL Data Library, November 1995
 A;Description: The sequence of C. elegans cosmid D2024.
 A;Reference number: Z21488
 A;Accession: T34200
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-435 <DUZ>
 A;Cross-references: EMBL:U41011; PIDN:AAA82288.1; CESP:D2024.3
 C;Genetics:
 A;Gene: CESP:D2024.3
 A;Introns: 17/3; 57/3; 306/1; 398/3

Query Match 100.0%; Score 31; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 215 NYGVH 219

RESULT 10
 S76734
 hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C;Species: Synechocystis sp.
 A;Variety: PCC 6803
 C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C;Accession: S76734
 R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.

A;Reference number: S74322; MUID:97061201; PMID:8905231
 A;Accession: S76734
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-475 <KAN>
 A;Cross-references: EMBL:D90916; GS:AB001339; NID:G1653715; PIDN:BAAL8646.1; PID:G1653733
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 100.0%; Score 31; DB 2; Length 475;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 |||||
 Db 347 NYGVH 351

RESULT 11
 E96500
 probable histidine decarboxylase [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C;Accession: E96500
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: E96500
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-462 <STO>
 A;Cross-references: GB:AE005173; NID:G7523682; PIDN:AAF63121.1; GSPDB:GN00141
 C;Genetics:

A:Gene: F2J6.7
A:Map position: 1
A:Superfamily: Klebsiella histidine decarboxylase

Query Match 100.0%; Score 31; DB 2; Length 482;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 139 NYGVH 143

RESULT 12

hypothetical protein C18E3.7 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15179
R:Connell, M.; Maggi, L.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of *C. elegans* cosmid C18E3.
A:Reference number: Z18304
A:Accession: T15179
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-697 <CON>
A:Cross-references: EMBL:AF000265; NID:g1947147; PID:g1947154; PIDN:AAB52947.1; GSPDB:GN
A:Experimental source: strain Bristol N2; clone C18E3
C:Genetics:
A:Map position: 1
A:Gene: CESP:C18E3.7
A:Introns: 12/2; 32/2; 72/1; 219/3; 315/2; 447/3; 666/2

Query Match 100.0%; Score 31; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 120 NYGVH 124

RESULT 13

S43529
165K protein, skeletal muscle - human
C:Species: *Homo sapiens* (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S43529; S42166
R:Fuerst, D.O.
submitted to the EMBL Data Library, October 1992
A:Reference number: S43529
A:Accession: S43529
A:Molecule type: mRNA
A:Residues: 1-1465 <FUE>
A:Cross-references: EMBL:X69089; NID:g407096; PIDN:CAA48832.1; PID:g407097
R:Vinkemeyer, U.; Obermann, W.; Weber, K.; Fuerst, D.O.
J. Cell Sci. 106, 319-330, 1993
A:Title: The globular head domain of titin extends into the center of the sarcomeric M band
A:Reference number: S42166; MUID:94095665; PMID:7505783
A:Accession: S42166
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-101, 'OR', 104-1465 <VIN>
A:Cross-references: EMBL:X69089
C:Superfamily: skeleton
C:Keywords: skeletal muscle

Query Match 100.0%; Score 31; DB 2; Length 1465;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5

Db 208 NYGVH 212

RESULT 14

T31076
tyrocidine synthetase 3 - *Brevibacillus brevis*
C:Species: *Brevibacillus brevis*
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
C:Accession: T31076
R:Mootz, H.D.; Marahiel, M.A.
J. Bacteriol. 179, 6843-6850, 1997
A:Title: The tyrocidine biosynthesis operon of *Bacillus brevis*: Complete nucleotide sequence
A:Reference number: Z20969; MUID:98012987; PMID:9352938
A:Accession: T31076
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6486 <MOO>
A:Cross-references: EMBL:AF004835; NID:g2623770; PID:g2623773; PIDN:AAC45930.1
C:Genetics:
A:Gene: tyoC
C:Function:
A:Pathway: tyrocidine biosynthesis
C:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:510-950/Domain: acetate-CoA ligase homology <ACLI>
F:968-1036/Domain: acyl carrier protein homology <ACP1>
F:1546-1987/Domain: acetate-CoA ligase homology <ACP2>
F:2005-2073/Domain: acyl carrier protein homology <ACP2>
F:2583-3025/Domain: acetate-CoA ligase homology <ACLI3>
F:3043-3111/Domain: acyl carrier protein homology <ACP3>
F:3621-4060/Domain: acetate-CoA ligase homology <ACLI4>
F:4078-4146/Domain: acyl carrier protein homology <ACP4>
F:4656-5104/Domain: acetate-CoA ligase homology <ACLI5>
F:5122-5190/Domain: acyl carrier protein homology <ACP5>
F:5702-6147/Domain: acetate-CoA ligase homology <ACLI6>
F:6165-6233/Domain: acyl carrier protein homology <ACP6>
F:1000-2037,3075,4110,5154,6197/Binding site: phosphopantetheine (Ser) (covalent) #status:

Query Match 100.0%; Score 31; DB 2; Length 6486;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 2664 NYGVH 2668

RESULT 15

PQ0266
Ig heavy chain V region (MCL) - mouse (fragment)
C:Species: *Mus musculus* (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-May-1997
C:Accession: PQ0266
R:Lochman, K.L.; Carrillo, M.A.; Kennedy, R.C.
Gene 105, 283-284, 1991
A:Title: Sequence analysis of the variable region of a mouse gene encoding a monoclonal antibody
A:Reference number: PQ0265; MUID:92039046; PMID:1937027
A:Accession: PQ0266
A:Molecule type: mRNA
A:Residues: 1-118 <LOH>
A:Cross-references: GB:M59984
C:Comment: This protein recognizes a restricted idiotype associated with antibodies specific for immunoglobulin V region; immunoglobulin homology
C:Superfamily: immunoglobulin V region; immunoglobulin
C:Keywords: heterotrimer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMV>
F:31-35/Region: complementarity-determining 1
F:50-65/Region: complementarity-determining 2
F:98-108/Region: complementarity-determining 3

Query Match 96.8%; Score 30; DB 2; Length 118;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
| | | |
Db 31 NYGIH 35

Search completed: October 6, 2004, 16:34:45
Job time : 8.26316 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 31.1404 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-2

Perfect score: 31

Sequence: 1 NYGVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------|
| 1 | 31 | 100.0 | 5 | 10 | US-09-798-689-26 |
| 2 | 31 | 100.0 | 5 | 10 | US-09-995-954B-2 |
| 3 | 31 | 100.0 | 5 | 12 | US-10-374-600-1 |
| 4 | 31 | 100.0 | 5 | 15 | US-10-374-531-1 |
| 5 | 31 | 100.0 | 19 | 10 | US-09-759-130B-231 |
| 6 | 31 | 100.0 | 19 | 16 | US-10-741-790-231 |
| 7 | 31 | 100.0 | 65 | 12 | US-10-424-599-193465 |
| 8 | 31 | 100.0 | 68 | 16 | US-10-767-701-50344 |
| 9 | 31 | 100.0 | 69 | 12 | US-10-424-599-187665 |
| 10 | 31 | 100.0 | 90 | 12 | US-10-424-599-218420 |
| 11 | 31 | 100.0 | 93 | 12 | US-10-424-599-2181738 |
| 12 | 31 | 100.0 | 116 | 12 | US-10-435-299-3 |
| 13 | 31 | 100.0 | 116 | 12 | US-10-435-299-4 |
| 14 | 31 | 100.0 | 119 | 12 | US-10-374-600-115 |
| 15 | 31 | 100.0 | 119 | 12 | US-10-374-600-116 |

| | | | | | | |
|----|----|-------|-----|----|-------------------|-------------------|
| 16 | 31 | 100.0 | 119 | 12 | US-10-374-600-117 | Sequence 117, App |
| 17 | 31 | 100.0 | 119 | 12 | US-10-374-600-118 | Sequence 118, App |
| 18 | 31 | 100.0 | 119 | 12 | US-10-374-600-119 | Sequence 119, App |
| 19 | 31 | 100.0 | 119 | 12 | US-10-374-600-120 | Sequence 120, App |
| 20 | 31 | 100.0 | 119 | 15 | US-10-374-531-115 | Sequence 115, App |
| 21 | 31 | 100.0 | 119 | 15 | US-10-374-531-116 | Sequence 116, App |
| 22 | 31 | 100.0 | 119 | 15 | US-10-374-531-117 | Sequence 117, App |
| 23 | 31 | 100.0 | 119 | 15 | US-10-374-531-118 | Sequence 118, App |
| 24 | 31 | 100.0 | 119 | 15 | US-10-374-531-119 | Sequence 119, App |
| 25 | 31 | 100.0 | 119 | 15 | US-10-374-531-120 | Sequence 120, App |
| 26 | 31 | 100.0 | 132 | 12 | US-10-374-600-24 | Sequence 24, Appl |
| 27 | 31 | 100.0 | 132 | 12 | US-10-374-600-26 | Sequence 26, Appl |
| 28 | 31 | 100.0 | 132 | 12 | US-10-374-600-27 | Sequence 27, Appl |
| 29 | 31 | 100.0 | 132 | 12 | US-10-374-600-28 | Sequence 28, Appl |
| 30 | 31 | 100.0 | 132 | 12 | US-10-374-600-29 | Sequence 29, Appl |
| 31 | 31 | 100.0 | 132 | 12 | US-10-374-600-30 | Sequence 30, Appl |
| 32 | 31 | 100.0 | 132 | 15 | US-10-374-531-24 | Sequence 24, Appl |
| 33 | 31 | 100.0 | 132 | 15 | US-10-374-531-26 | Sequence 26, Appl |
| 34 | 31 | 100.0 | 132 | 15 | US-10-374-531-27 | Sequence 27, Appl |
| 35 | 31 | 100.0 | 132 | 15 | US-10-374-531-28 | Sequence 28, Appl |
| 36 | 31 | 100.0 | 132 | 15 | US-10-374-531-29 | Sequence 29, Appl |
| 37 | 31 | 100.0 | 132 | 15 | US-10-374-531-30 | Sequence 30, Appl |
| 38 | 31 | 100.0 | 138 | 12 | US-10-374-600-8 | Sequence 8, Appl |
| 39 | 31 | 100.0 | 138 | 12 | US-10-374-600-13 | Sequence 13, Appl |
| 40 | 31 | 100.0 | 138 | 12 | US-10-374-600-19 | Sequence 19, Appl |
| 41 | 31 | 100.0 | 138 | 15 | US-10-374-531-8 | Sequence 8, Appl |
| 42 | 31 | 100.0 | 138 | 15 | US-10-374-531-13 | Sequence 13, Appl |
| 43 | 31 | 100.0 | 138 | 15 | US-10-374-531-19 | Sequence 19, Appl |
| 44 | 31 | 100.0 | 141 | 9 | US-09-903-456-47 | Sequence 47, Appl |
| 45 | 31 | 100.0 | 141 | 14 | US-10-456-911-47 | Sequence 47, Appl |

ALIGNMENTS

RESULT 1
US-09-798-689-26
; Sequence 26, Application US/09798689
; Publication No. US20030103973A1
; GENERAL INFORMATION:
; APPLICANT: Rockwell, Patricia
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: Method for Reducing Tumor Growth with VEGF Antagonists
; TITLE OF INVENTION: Combined With Radiation and Chemotherapy
; FILE REFERENCE: Sequence Listings 1-41 for 381-25 CIP
; CURRENT APPLICATION NUMBER: US/09798,689
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/401,163
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/967,113
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 08/706,804
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/476,533
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/326,552
; PRIOR FILING DATE: 1994-10-20
; PRIOR APPLICATION NUMBER: 08/196,041
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26

Query Match 100.0%; Score 31; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
|||||

Db 1 NYGVH 5

RESULT 2

US-09-996-954B-2

Sequence 1, Application US/09996954B

Publication No. US20030157104A1

GENERAL INFORMATION:

APPLICANT: Waksal, Harlan W.

TITLE OF INVENTION: Treatment of Refractory Human Tumors

TITLE OF INVENTION: With Epidermal Growth Factor Receptor Antagonists

FILE REFERENCE: 11245-46605

CURRENT APPLICATION NUMBER: US/09/996,954B

CURRENT FILING DATE: 2001-11-30

PRIOR APPLICATION NUMBER: 09/840,146

PRIOR FILING DATE: 04-24-2001

PRIOR APPLICATION NUMBER: 09/374,028

PRIOR FILING DATE: 08-13-1999

PRIOR APPLICATION NUMBER: 09/312,284

PRIOR FILING DATE: 05-14-1999

NUMBER OF SEQ ID NOS: 12

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 5

TYPE: PRT

ORGANISM: Homo sapiens-Rodent Chimera

US-09-996-954B-2

Query Match 100.0%; Score 31; DB 10; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 1 NYGVH 5

RESULT 3

US-10-374-600-1

Sequence 1, Application US/10374600

Publication No. US20030224001A1

GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.

TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,600

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-374-600-1

Query Match 100.0%; Score 31; DB 12; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5

Db 1 NYGVH 5

RESULT 4

US-10-374-531-1

Sequence 1, Application US/10374531

Publication No. US20040006212A1

GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.

TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,531

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

```

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-374-531-1

Query Match      100.0%; Score 31; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NYGVH 5
Db      1 NYGVH 5

RESULT 5
US-09-759-130B-231
; Sequence 231, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-5350NMIM
; CURRENT APPLICATION NUMBER: US 09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 231
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-231

Query Match      100.0%; Score 31; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NYGVH 5
Db      6 NYGVH 10

US-09-759-130B-231

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-374-531-1

Query Match      100.0%; Score 31; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NYGVH 5
Db      1 NYGVH 5

RESULT 6
US-10-741-790-231
; Sequence 231, Application US/10741790
; Publication No. US20040121396A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-5350NMIM
; CURRENT APPLICATION NUMBER: US/10741,790
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 231
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-790-231

Query Match      100.0%; Score 31; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 NYGVH 5
Db      6 NYGVH 10

RESULT 7
US-10-424-599-193465
; Sequence 193465, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 193465
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_16723C.1.pep
US-10-424-599-193465

Query Match      100.0%; Score 31; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 41 NYGVH 45

RESULT 8
US-10-767-701-50344
; Sequence 50344, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50344
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-058-Q6-K1-A9.pep
US-10-767-701-50344

Query Match      100.0%; Score 31; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 10 NYGVH 14

RESULT 9
US-10-424-599-187665
; Sequence 187665, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187665
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140475C.1.pep

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 193465
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_16723C.1.pep
US-10-424-599-193465

Query Match      100.0%; Score 31; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 41 NYGVH 45

RESULT 8
US-10-767-701-50344
; Sequence 50344, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 50344
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3478-058-Q6-K1-A9.pep
US-10-767-701-50344

Query Match      100.0%; Score 31; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 10 NYGVH 14

RESULT 9
US-10-424-599-187665
; Sequence 187665, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187665
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140475C.1.pep

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 281738
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(93)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9642C.1.pep
US-10-424-599-281738

Query Match      100.0%; Score 31; DB 12; Length 93;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 NYGVH 5
 Db 14 NYGVH 18

RESULT 12

US-10-435-299-3
 ; Sequence 3, Application US/10435299
 ; Publication No. US20040052783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weiner, George
 ; APPLICANT: Gingrich, Roger
 ; APPLICANT: Link, Brian
 ; APPLICANT: Tso, J. Yun
 ; TITLE OF INVENTION: HUMANIZED ANTIBODIES AGAINST CD3
 ; FILE REFERENCE: 05882-0176-CNUS04
 ; CURRENT APPLICATION NUMBER: US/10/435,299
 ; CURRENT FILING DATE: 2003-05-09
 ; PRIOR APPLICATION NUMBER: US 09/618,380
 ; PRIOR FILING DATE: 2000-07-18
 ; PRIOR APPLICATION NUMBER: US 08/397,411
 ; PRIOR FILING DATE: 1995-03-01
 ; PRIOR APPLICATION NUMBER: US 07/859,583
 ; PRIOR FILING DATE: 1992-03-27
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 116
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Heavy chain of Humanized ID10 Ab minus signal sequence
 US-10-435-299-3

Query Match 100.0%; Score 31; DB 12; Length 116;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 13

US-10-435-299-4
 ; Sequence 4, Application US/10435299
 ; Publication No. US20040052783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weiner, George
 ; APPLICANT: Gingrich, Roger
 ; APPLICANT: Link, Brian
 ; APPLICANT: Tso, J. Yun
 ; TITLE OF INVENTION: HUMANIZED ANTIBODIES AGAINST CD3
 ; FILE REFERENCE: 05882-0176-CNUS04
 ; CURRENT APPLICATION NUMBER: US/10/435,299
 ; CURRENT FILING DATE: 2003-05-09
 ; PRIOR APPLICATION NUMBER: US 09/618,380
 ; PRIOR FILING DATE: 2000-07-18
 ; PRIOR APPLICATION NUMBER: US 08/397,411
 ; PRIOR FILING DATE: 1995-03-01
 ; PRIOR APPLICATION NUMBER: US 07/859,583
 ; PRIOR FILING DATE: 1992-03-27
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 116
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 US-10-435-299-4

Query Match 100.0%; Score 31; DB 12; Length 116;
 Best Local Similarity 100.0%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 14

US-10-374-600-115
 ; Sequence 115, Application US/10374600
 ; Publication No. US20030224001A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ImClone Systems Incorporated, et al.
 ; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
 ; INHIBITING THE GROWTH OF TUMORS
 ; NUMBER OF SEQUENCES: 120
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenyon & Kenyon
 ; STREET: One Broadway
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: US
 ; ZIP: 10004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/374,600
 ; FILING DATE: 25-Feb-2003
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/973,065C
 ; FILING DATE: 19-Mar-1998
 ; APPLICATION NUMBER: PCT/US96/09847
 ; FILING DATE: 07-JUN-1996
 ; APPLICATION NUMBER: US 08/482,982
 ; FILING DATE: 07-JUN-1995
 ; APPLICATION NUMBER: US 08/573,289
 ; FILING DATE: 15-DEC-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Deborah A. Somerville
 ; REGISTRATION NUMBER: 31,995
 ; REFERENCE/DOCKET NUMBER: 11245/46003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 425-7200
 ; TELEFAX: (212) 425-5288
 ; INFORMATION FOR SEQ ID NO: 115:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 119 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: peptide
 ; HYPOTHEetical: NO
 ; FRAGMENT TYPE: internal
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 115:
 US-10-374-600-115

Query Match 100.0%; Score 31; DB 12; Length 119;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 15

US-10-374-600-116
 ; Sequence 116, Application US/10374600
 ; Publication No. US20030224001A1

```

; GENERAL INFORMATION:
; APPLICANT: InClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-10-374-600-116

```

```

Query Match      100.0%; Score 31; DB 12; Length 119;
Best Local Similarity 100.0%; Pred.No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 NYGVH 5
Db      31 NYGVH 35

```

```

Search completed: October 6, 2004, 17:09:02
Job time : 34.1404 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 8.85965 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-2

Perfect score: 31

Sequence: 1 NYGVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
 - 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
 - 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
 - 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
 - 5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
 - 6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 31 | 100.0 | 116 | 3 | US-08-397-411-3 |
| 2 | 31 | 100.0 | 116 | 3 | US-08-397-411-4 |
| 3 | 31 | 100.0 | 119 | 3 | US-08-483-749A-14 |
| 4 | 31 | 100.0 | 141 | 4 | US-09-903-456-47 |
| 5 | 31 | 100.0 | 147 | 4 | US-09-145-828A-20 |
| 6 | 31 | 100.0 | 147 | 4 | US-09-903-456-27 |
| 7 | 31 | 100.0 | 239 | 6 | 5455030-13 |
| 8 | 31 | 100.0 | 242 | 6 | 5455030-15 |
| 9 | 31 | 100.0 | 265 | 4 | US-09-903-456-64 |
| 10 | 31 | 100.0 | 273 | 3 | US-08-397-411-6 |
| 11 | 31 | 100.0 | 446 | 3 | US-08-397-411-7 |
| 12 | 30 | 96.8 | 505 | 4 | US-09-134-000C-5340 |
| 13 | 30 | 96.8 | 680 | 1 | US-08-211-430-2 |
| 14 | 30 | 96.8 | 680 | 3 | US-08-761-136-1 |
| 15 | 30 | 96.8 | 680 | 4 | US-09-576-967-1 |
| 16 | 28 | 90.3 | 120 | 2 | US-08-652-558-39 |
| 17 | 28 | 90.3 | 125 | 3 | US-09-240-274-8 |
| 18 | 28 | 90.3 | 125 | 3 | US-09-240-274-12 |
| 19 | 28 | 90.3 | 125 | 3 | US-09-240-274-20 |
| 20 | 28 | 90.3 | 125 | 3 | US-09-240-274-21 |
| 21 | 28 | 90.3 | 125 | 3 | US-09-240-274-22 |
| 22 | 28 | 90.3 | 125 | 3 | US-09-240-274-23 |
| 23 | 28 | 90.3 | 126 | 3 | US-09-240-274-14 |
| 24 | 28 | 90.3 | 126 | 3 | US-09-240-274-15 |
| 25 | 28 | 90.3 | 126 | 3 | US-09-240-274-16 |
| 26 | 28 | 90.3 | 126 | 3 | US-09-240-274-17 |
| 27 | 28 | 90.3 | 126 | 3 | US-09-240-274-148 |

| | | | | | | |
|----|----|------|-----|---|---------------------|-------------------|
| 28 | 28 | 90.3 | 166 | 4 | US-09-472-087-75 | Sequence 75, Appl |
| 29 | 28 | 90.3 | 187 | 4 | US-09-328-352-4950 | Sequence 4950, Ap |
| 30 | 28 | 90.3 | 225 | 4 | US-09-456-030A-68 | Sequence 68, Appl |
| 31 | 28 | 90.3 | 233 | 2 | US-08-480-753-2 | Sequence 2, Appl |
| 32 | 28 | 90.3 | 233 | 3 | US-09-041-889-9 | Sequence 9, Appl |
| 33 | 28 | 90.3 | 233 | 3 | US-08-837-058-9 | Sequence 9, Appl |
| 34 | 28 | 90.3 | 233 | 4 | US-09-417-264-9 | Sequence 9, Appl |
| 35 | 28 | 90.3 | 244 | 2 | US-08-480-753-4 | Sequence 4, Appl |
| 36 | 28 | 90.3 | 347 | 4 | US-09-134-001C-3752 | Sequence 3752, Ap |
| 37 | 28 | 90.3 | 464 | 4 | US-09-472-087-2 | Sequence 2, Appl |
| 38 | 28 | 90.3 | 464 | 4 | US-08-472-087-66 | Sequence 66, Appl |
| 39 | 28 | 90.3 | 498 | 4 | US-09-489-039A-7985 | Sequence 7985, Ap |
| 40 | 28 | 90.3 | 701 | 4 | US-09-328-352-5717 | Sequence 5717, Ap |
| 41 | 27 | 87.1 | 228 | 4 | US-09-724-864-42 | Sequence 42, Appl |
| 42 | 27 | 87.1 | 271 | 4 | US-09-903-456-53 | Sequence 53, Appl |
| 43 | 27 | 87.1 | 272 | 4 | US-09-145-828A-18 | Sequence 18, Appl |
| 44 | 27 | 87.1 | 272 | 4 | US-09-903-456-24 | Sequence 24, Appl |
| 45 | 27 | 87.1 | 319 | 4 | US-09-134-000C-5242 | Sequence 5242, Ap |

ALIGNMENTS

RESULT 1
US-08-397-411-3
; Sequence 3, Application US/08397411
; Patent No. 6129914
; GENERAL INFORMATION:
; APPLICANT: Weiner, George
; APPLICANT: Gingrich, Roger
; APPLICANT: Link, Brian
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: Bispecific Antibody Effective to Treat
; TITLE OF INVENTION: B-Cell Lymphoma and Cell Line
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/397,411
; FILING DATE: 01-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/859,583
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 011823-004901
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 116 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-411-3

Query Match 100.0%; Score 31; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 2

US-08-397-411-4
 ; Sequence 4, Application US/08397411
 ; Patent No. 6129914
 ; GENERAL INFORMATION:
 ; APPLICANT: Weiner, George
 ; APPLICANT: Gingrich, Roger
 ; APPLICANT: Link, Brian
 ; APPLICANT: Tso, J. Yun
 ; TITLE OF INVENTION: Bispecific Antibody Effective to Treat
 ; TITLE OF INVENTION: B-Cell Lymphoma and Cell Line
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Townsend and Townsend and Crew
 ; STREET: One Market Plaza, Steuart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/397,411
 ; FILING DATE: 01-MAR-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/859,583
 ; FILING DATE: 27-MAR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, William M.
 ; REGISTRATION NUMBER: 30,223
 ; REFERENCE/DOCKET NUMBER: 011823-004901
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-326-2400
 ; TELEFAX: 415-326-2422
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 116 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-397-411-4

Query Match 100.0%; Score 31; DB 3; Length 116;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 3

US-08-483-749A-14
 ; Sequence 14, Application US/08483749A
 ; Patent No. 6054561
 ; GENERAL INFORMATION:
 ; APPLICANT: RING, DAVID B.
 ; TITLE OF INVENTION: ANTIGEN-BINDING SITES OF ANTIBODY
 ; TITLE OF INVENTION: MOLECULES SPECIFIC FOR CANCER ANTIGENS
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CHIRON CORPORATION

STREET: INTELLECTUAL PROPERTY - R440, PO BOX 8097
 CITY: EMERYVILLE
 STATE: CA
 COUNTRY: USA
 ZIP: 94662-8097
 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483,749A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: SAVERIDE, PAUL B.
 ; REGISTRATION NUMBER: 36,914
 ; REFERENCE/DOCKET NUMBER: 0508.008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 601-2585
 ; TELEFAX: (510) 655-3542
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 119 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-483-749A-14

Query Match 100.0%; Score 31; DB 3; Length 119;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
 Db 31 NYGVH 35

RESULT 4

US-09-903-456-47
 ; Sequence 47, Application US/09903456
 ; Patent No. 6677145
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott Laboratories
 ; APPLICANT: Mukerji, Pradip
 ; APPLICANT: Leonard, Amanda Eun-Yeong
 ; APPLICANT: Huang, Yung-Sheng
 ; APPLICANT: Pereira, Suzette L.
 ; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
 ; FILE REFERENCE: 6407.US.P3
 ; CURRENT APPLICATION NUMBER: US/09/903,456
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 09/624,670
 ; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: US 09/379,095
 ; PRIOR FILING DATE: 1999-08-23
 ; PRIOR APPLICATION NUMBER: US 09/145,828
 ; PRIOR FILING DATE: 1998-09-02
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 141
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (141)...(141)
 ; OTHER INFORMATION: Xaa = Unknown or Other at position 141
 ; US-09-903-456-47

Query Match 100.0%; Score 31; DB 4; Length 141;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 45 NYGVH 49

RESULT 5
US-09-145-828A-20
; Sequence 20, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407 US 01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-145-828A-20

Query Match 100.0%; Score 31; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 63 NYGVH 67

RESULT 6
US-09-903-456-27
; Sequence 27, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Perelra, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-903-456-27

Query Match 100.0%; Score 31; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 63 NYGVH 67

RESULT 7
5455030-13

; Patent No. 5455030
; APPLICANT: LADNER, ROBERT C.; BIRD, ROBERT E.; HARDMAN, KARL
; TITLE OF INVENTION: IMMUNOTHERAPY USING SINGLE CHAIN
; POLYPEPTIDE BINDING MOLECULES
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/40,440
; FILING DATE: 1-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 512,910
; FILING DATE: 25-APR-1990
; APPLICATION NUMBER: 299,617
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 92,110
; FILING DATE: 02-SEP-1987
; APPLICATION NUMBER: 902,971
; FILING DATE: 01-SEP-1986
; SEQ ID NO: 13
; LENGTH: 239
5455030-13

Query Match 100.0%; Score 31; DB 6; Length 239;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 151 NYGVH 155

RESULT 8
5455030-15
; Patent No. 5455030
; APPLICANT: LADNER, ROBERT C.; BIRD, ROBERT E.; HARDMAN, KARL
; TITLE OF INVENTION: IMMUNOTHERAPY USING SINGLE CHAIN
; POLYPEPTIDE BINDING MOLECULES
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/40,440
; FILING DATE: 1-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 512,910
; FILING DATE: 25-APR-1990
; APPLICATION NUMBER: 299,617
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 92,110
; FILING DATE: 02-SEP-1987
; APPLICATION NUMBER: 902,971
; FILING DATE: 01-SEP-1986
; SEQ ID NO: 15
; LENGTH: 242
5455030-15

Query Match 100.0%; Score 31; DB 6; Length 242;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYGVH 5
Db 153 NYGVH 157

RESULT 9
US-09-903-456-64
; Sequence 64, Application US/09903456
; Patent No. 6677145
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories

APPLICANT: Mukerji, Pradip
APPLICANT: Leonard, Amanda Eun-Yeong
APPLICANT: Huang, Yung-Sheng
APPLICANT: Persira, Suzette L.
TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
FILE REFERENCE: 6407.US.P3
CURRENT APPLICATION NUMBER: US/09/903,456
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: US 09/624,670
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: US 09/379,095
PRIOR FILING DATE: 1999-08-23
PRIOR APPLICATION NUMBER: US 09/145,828
PRIOR FILING DATE: 1998-09-02
NUMBER OF SEQ ID NOS: 116
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 64
LENGTH: 265
TYPE: PRT
ORGANISM: Homo sapiens
US-09-903-456-64

Query Match 100.0%; Score 31; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 170 NYGVH 174

RESULT 10
US-08-397-411-6
Sequence 6, Application US/08397411
Patent No. 6129914
GENERAL INFORMATION:
APPLICANT: Weiner, George
APPLICANT: Gingrich, Roger
APPLICANT: Link, Brian
APPLICANT: Tso, J. Yun
TITLE OF INVENTION: Bispecific Antibody Effective to Treat
TITLE OF INVENTION: B-Cell Lymphoma and Cell Line
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/397,411
FILING DATE: 01-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/859,583
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 011823-004901
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-397-411-6
Query Match 100.0%; Score 31; DB 3; Length 273;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 31 NYGVH 35

RESULT 11
US-08-397-411-7
Sequence 7, Application US/08397411
Patent No. 6129914
GENERAL INFORMATION:
APPLICANT: Weiner, George
APPLICANT: Gingrich, Roger
APPLICANT: Link, Brian
APPLICANT: Tso, J. Yun
TITLE OF INVENTION: Bispecific Antibody Effective to Treat
TITLE OF INVENTION: B-Cell Lymphoma and Cell Line
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/397,411
FILING DATE: 01-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/859,583
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 011823-004901
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 446 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-397-411-7

Query Match 100.0%; Score 31; DB 3; Length 446;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 31 NYGVH 35

RESULT 12
US-09-134-000C-5340
Sequence 5340, Application US/09134000C

```

; Patent No. 5617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5340
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-08-211-430-2

Query Match          96.8%; Score 30; DB 4; Length 505;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 195 NYGIH 199

RESULT 13
US-08-211-430-2
; Sequence 2, Application US/08211430
; Patent No. 5763166
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC SEQUENCE OF THE GENE ASSOCIATED WITH
; TITLE OF INVENTION: X CHROMOSOME LINKED KALLMANN SYNDROME, CORRESPONDING
; NUMBER OF SEQUENCES: 32
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION FOR SEQ ID NO: 2:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 680 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: foetal brain cell
; US-08-211-430-2

Query Match          96.8%; Score 30; DB 1; Length 680;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 224 NYGIH 228

RESULT 14
US-08-761-136-1
; Sequence 1, Application US/08761136
; Patent No. 6121231
; GENERAL INFORMATION:
; APPLICANT: PETIT, CHRISTINE

```

```

; APPLICANT: SOUSSI-YANTICOCTAS, NADIA
; APPLICANT: HARDELIN, JEAN-PIERRE
; APPLICANT: SARAILH, CATHERINE
; APPLICANT: ROUGON, GENEVIEVE
; APPLICANT: LEGOUIS, RENAUD
; APPLICANT: ARDOUIN, OLIVIER
; APPLICANT: MAZIE, JEAN-CLAUDE
; TITLE OF INVENTION: USE OF KAL PROTEIN AND TREATMENT WITH
; TITLE OF INVENTION: THE KAL PROTEIN IN TREATMENT OF RETINAL, RENAL, NEUROMAL
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,136
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 650-112-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-412-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 680 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-761-136-1

Query Match          96.8%; Score 30; DB 3; Length 680;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
Db 224 NYGIH 228

RESULT 15
US-09-576-967-1
; Sequence 1, Application US/09576967
; Patent No. 6548475
; GENERAL INFORMATION:
; APPLICANT: PETIT, CHRISTINE
; APPLICANT: SOUSSI-YANTICOCTAS, NADIA
; APPLICANT: HARDELIN, JEAN-PIERRE
; APPLICANT: SARAILH, CATHERINE
; APPLICANT: ROUGON, GENEVIEVE
; APPLICANT: LEGOUIS, RENAUD
; APPLICANT: ARDOUIN, OLIVIER
; APPLICANT: MAZIE, JEAN-CLAUDE
; TITLE OF INVENTION: USE OF KAL PROTEIN AND TREATMENT WITH
; TITLE OF INVENTION: THE KAL PROTEIN IN TREATMENT OF RETINAL, RENAL, NEUROMAL
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.

```

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,967
FILING DATE: 24-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/761,136
FILING DATE: 06-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-112-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-412-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 680 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-576-967-1

Query Match 96.8%; Score 30; DB 4; Length 680;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYGVH 5
DB 224 NYGIH 228

Search completed: October 6, 2004, 16:36:30
Job time : 9.85965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:59 ; Search time 6.31579 Seconds
(without alignments)
91.382 Million cell updates/sec

Title: US-09-635-974A-10

Perfect score: 25

Sequence: 1 ASESIS 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 25 | 100.0 | 87 | 2 PH1082 | Ig light chain V r |
| 2 | 25 | 100.0 | 92 | 1 KVRB38 | Ig kappa chain V r |
| 3 | 25 | 100.0 | 108 | 2 C30502 | Ig kappa chain V r |
| 4 | 25 | 100.0 | 115 | 1 KWSL7 | Ig kappa chain pre |
| 5 | 25 | 100.0 | 128 | 2 PNO445 | Ig kappa chain pre |
| 6 | 25 | 100.0 | 248 | 2 E90552 | hypothetical prote |
| 7 | 25 | 100.0 | 368 | 2 S17980 | pgal protein - Asp |
| 8 | 25 | 100.0 | 649 | 2 T37740 | coiled-coil protei |
| 9 | 25 | 100.0 | 732 | 2 T50143 | pumilio family pro |
| 10 | 25 | 100.0 | 1188 | 2 JCA4889 | phosphatidylinosit |
| 11 | 25 | 100.0 | 1392 | 2 A43336 | microtubule-vesicl |
| 12 | 25 | 100.0 | 1427 | 2 S22695 | restin - human |
| 13 | 24 | 96.0 | 304 | 2 F84169 | hypothetical prote |
| 14 | 24 | 96.0 | 349 | 2 T19485 | hypothetical prote |
| 15 | 24 | 96.0 | 443 | 2 T13299 | probable helicase |
| 16 | 24 | 96.0 | 491 | 2 S49790 | probable membrane |
| 17 | 24 | 96.0 | 635 | 2 A81656 | threonyl-tRNA synt |
| 18 | 24 | 96.0 | 683 | 2 JCS393 | zinc finger protei |
| 19 | 24 | 96.0 | 685 | 2 JCS392 | zinc finger protei |
| 20 | 24 | 96.0 | 691 | 2 T48506 | hypothetical prote |
| 21 | 24 | 96.0 | 757 | 2 S68142 | probable transcrip |
| 22 | 24 | 96.0 | 893 | 2 A47550 | bride of sevenless |
| 23 | 24 | 96.0 | 1532 | 2 A26039 | IGA-specific metal |
| 24 | 24 | 96.0 | 1639 | 2 T50119 | probable sensory t |
| 25 | 24 | 96.0 | 2570 | 2 T17451 | fibriae-associate |
| 26 | 24 | 96.0 | 3670 | 2 T36249 | CDA peptide synth |
| 27 | 23 | 92.0 | 316 | 2 B71569 | hypothetical prote |
| 28 | 23 | 92.0 | 340 | 2 F82468 | hypothetical prote |
| 29 | 23 | 92.0 | 511 | 2 AB0858 | hypothetical prote |

30 23 92.0 583 2 S67571 hypothetical prote
31 23 92.0 584 2 G56713 hypothetical prote
32 23 616 2 A50824 chaperone protein
33 23 92.0 711 2 A37051 outer membrane pro
34 23 92.0 720 2 S70838 pilQ protein - Nei
35 23 92.0 761 2 A81985 pilus secretin NWA
36 23 92.0 1473 2 T13855 suppressor of sabl
37 23 92.0 1539 2 T30037 hypothetical prote
38 23 92.0 2251 2 T24490 hypothetical prote
39 22 88.0 69 2 PH1080 Ig light chain V r
40 22 88.0 86 2 S78488 Ig kappa chain V r
41 22 88.0 86 2 S16837 Ig kappa chain V r
42 22 88.0 86 2 S34086 Ig kappa chain V r
43 22 88.0 86 2 S16840 Ig kappa chain V r
44 22 88.0 87 2 S34084 Ig kappa chain V r
45 22 88.0 87 2 S34083 Ig kappa chain V r

ALIGNMENTS

RESULT 1

PH1082

Ig light chain V region (clone 165.54) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000

C:Accession: PH1082

R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.

J. Exp. Med. 176, 761-772, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH1082

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-87 <TIL>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin

F:6-80/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 25; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6

Db 41 ASESIS 46

RESULT 2

KVRB38

Ig kappa chain V region (3381) - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 21-Jan-2000

C:Accession: A01953

R:Margolies, M.N.; Cannon III, L.E.; Strosberg, A.D.; Haber, E.

Proc. Natl. Acad. Sci. U.S.A. 72, 2180-2184, 1975

A:Title: Diversity of light chain variable region sequences among rabbit antibodies elic

A:Reference number: A93799; MUID:75176905; PMID:1034456

A:Accession: A01953

A:Molecule type: protein

A:Residues: 1-92 <MAR>

C:Comment: This chain was obtained from antibody to type III pneumococci and was isolate

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer

F:16-90/Domain: immunoglobulin homology <IMM>

F:23-88/Disulfide bonds: #status predicted

Query Match 100.0%; Score 25; DB 1; Length 92;

Best Local Similarity 100.0%; Pred. No. 8.7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 Db 25 ASESIS 30

RESULT 3

Ig kappa chain V region (D444) - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Nov-1988 #sequence_revision 03-Nov-1988 #text_change 21-Jan-2000
 C/Accession: C30502
 R/Bilat, D.; Webster, D.M.; Rees, A.R.
 J. Immunol. 141, 1745-1753, 1988
 A/Title: V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F-1 mice
 A/Reference number: A30502; MUID:189315787; PMID:2457627
 A/Accession: C30502
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-108 <EIL>
 A/Cross-references: GB:M21907; NID:G197071; PIDN:AAA38907.1; PID:G197072
 C/Suprafamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotetramer; immunoglobulin
 F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 25; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 Db 51 ASESIS 56

RESULT 4

Ig kappa chain precursor V region (L7) - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 22-Jun-1999
 C/Accession: A01925
 R/Pech, M.; Hochl, J.; Schnell, H.; Zachau, H.G.
 Nature 291, 668-670, 1981
 A/Title: Differences between germ-line and rearranged immunoglobulin V-kappa coding sequences
 A/Reference number: A93259; MUID:181220975; PMID:6264318
 A/Accession: A01925
 A/Molecule type: DNA
 A/Residues: 1-115 <PEC>
 A/Cross-references: GB:V01564; GB:J00574; NID:G51718; PIDN:CAA24884.1; PID:G758153
 A/Note: the sequence was determined from the germline gene
 A/Note: there appear to be two possible splice junctions at the 3' end of the intron; the
 C/Genetics:
 A/Introns: 17/1
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger
 C/Suprafamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotetramer
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-115/Product: Ig kappa chain V region (L7) #status predicted <MAT>
 F:136-110/Domain: immunoglobulin homology <IMM>
 F:43-108/Disulfide bonds: #status predicted

Query Match 100.0%; Score 25; DB 1; Length 115;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 Db 71 ASESIS 76

RESULT 5

PN0445
 Ig kappa chain precursor V-I region - human (fragment)

C/Species: Homo sapiens (man)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
 C/Accession: PN0445
 R/Kaluza, B.; Betzl, G.; Shao, H.; Diamantstein, T.; Weidle, U.H.
 Gene 122, 321-328, 1992
 A/Title: A general method for chimerization of monoclonal antibodies by inverse polymerase
 A/Reference number: PN0444; MUID:93138402; PMID:1339379
 A/Accession: PN0445
 A/Molecule type: mRNA
 A/Residues: 1-128 <KAL>
 A/Cross-references: GB:I02347
 C/Suprafamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotetramer; immunoglobulin
 F:1-10/Domain: signal sequence #status predicted <SIG>
 F:11-128/Product: Ig light chain kappa-1 V region #status predicted <MAT>
 F:26-100/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 25; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 Db 61 ASESIS 66

RESULT 6

E90552
 hypothetical protein MYPV 3250 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
 C/Species: Mycoplasma pulmonis
 C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C/Accession: E90552
 R/Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
 A/Reference number: A99512; MUID:21267165; PMID:11353084
 A/Accession: E90552
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-248 <KUR>
 A/Cross-references: GB:AL445566; PID:G14089739; PIDN:CAC13498.1; GSPDB:GN00153
 A/Experimental source: strain UAB CTIP
 C/Genetics:
 A/Gene: MYPV_3250
 A/Genetic code: SGC3

Query Match 100.0%; Score 25; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 Db 11 ASESIS 16

RESULT 7

pgal protein - Aspergillus niger
 C/Species: Aspergillus niger
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
 C/Accession: S17980
 R/Bussink, H.J.D.; Brouwer, K.B.; de Graaff, L.H.; Kester, H.C.M.; Visser, J.
 Curr. Genet. 20, 301-307, 1991
 A/Title: Identification and characterization of a second polygalacturonase gene of Aspergillus
 A/Reference number: S17980; MUID:92035101; PMID:1934135
 A/Accession: S17980
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-368 <BUS>
 A/Cross-references: EMBL:X58892; NID:G2382; PIDN:CAA41693.1; PID:G2383
 C/Genetics:
 A/Introns: 76/3; 216/3
 C/Suprafamily: polygalacturonase

Qy 1 ASESIS 6
 Db 71 ASESIS 76

Query Match 100.0%; Score 25; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 |||||
 Db 43 ASESIS 48

RESULT 8
 T37740
 coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37740
 R:Rieger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999
 A:Reference number: Z21743
 A:Accession: T37740
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-649 <RIE>
 A:Cross-references: EMBL:AL035248; PIDN:CAA22848.1; GSPDB:GN00066; SPDB:SPAC167.03c
 A:Experimental source: strain 972h-; cosmid c167
 C:Genetics:
 A:Gene: SPDB:SPAC167.03c
 A:Map position: 1

Query Match 100.0%; Score 25; DB 2; Length 649;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 |||||
 Db 7 ASESIS 12

RESULT 9
 T50143
 pumilio family protein [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
 C:Accession: T50143
 R:Hamlin, N.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, November 1999
 A:Reference number: Z25043
 A:Accession: T50143
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-732 <HAM>
 A:Cross-references: EMBL:AL132798; PIDN:CAB60694.1; GSPDB:GN00066; SPDB:SPAC222.02c
 A:Experimental source: strain 972h(-); cosmid c222
 C:Genetics:
 A:Gene: SPAC1687.22c; SPDB:SPAC222.02c
 A:Map position: 1
 A:Introns: 656/2; 711/1

Query Match 100.0%; Score 25; DB 2; Length 732;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 |||||
 Db 11 ASESIS 16

RESULT 10
 JC4889
 phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase (EC 3.1.3.-) - human
 N:Alternate names: hpf5/CN5phosphatase
 C:Species: Homo sapiens (man)
 C>Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999

C:Accession: JC4889; PC4187
 R:Drayer, A.L.; Pesesse, X.; De Smedt, F.; Woscholski, R.; Parker, P.; Erneux, C.
 Biochem. Biophys. Res. Commun. 225, 243-249, 1996
 A:Title: Cloning and expression of a human placenta inositol 1,3,4,5-tetrakisphosphate,
 A:Reference number: JC4889; MUID:96332436; PMID:8769125
 A:Contents: placenta
 A:Accession: JC4889
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1188 <DRA>
 A:Cross-references: EMBL:X98429; NID:G1495455; PIDN:CAA67071.1; PID:e249440; PID:G14954
 A:Accession: PC4187
 A:Residues: 582-592; 668-675 <DR2>
 C:Comment: This enzyme plays a function in the phosphatidylinositol 3-kinase and isosit.
 sequence motifs show that this enzyme interacts with various proteins in signal transd.
 C:Comment: inositol 1,3,4,5-tetrakisphosphate is also a substrate for this enzyme.
 C:Superfamily: SH2 homology
 C:Keywords: phosphoric monoester hydrolase
 F:5-101/Domain: SH2 homology <SH2>
 F:380/Binding site: substrate (Arg) #status predicted
 F:671/Active site: Cys #status predicted

Query Match 100.0%; Score 25; DB 2; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 |||||
 Db 32 ASESIS 37

RESULT 11
 A43336
 microtubule-vesicle linker CLIP-170 - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
 C:Accession: A43336
 R:Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
 Cell 70, 887-900, 1992
 A:Title: CLIP-170 links endocytic vesicles to microtubules.
 A:Reference number: A43336; MUID:92405160; PMID:1356075
 A:Accession: A43336
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1392 <PIE>
 A:Cross-references: GB:M97501; NID:G180621; PIDN:AAA35693.1; PID:G180622

Query Match 100.0%; Score 25; DB 2; Length 1392;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
 |||||
 Db 192 ASESIS 197

RESULT 12
 S22695
 restin - Homo sapiens (man)
 C:Species: Homo sapiens (man)
 C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999
 C:Accession: S22695; S19953
 R:Bilber, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
 EMBO J. 11, 2103-2113, 1992
 A:Title: Restin: a novel intermediate filament-associated protein highly expressed in th
 A:Reference number: S22695; MUID:92289675; PMID:1600942
 A:Accession: S22695
 A:Molecule type: mRNA
 A:Residues: 1-1427 <BIL>
 A:Cross-references: EMBL:X64838; NID:G35998; PIDN:CAA46050.1; PID:G35999
 C:Keywords: cytoskeleton

Query Match 100.0%; Score 25; DB 2; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 |||||
 DB 192 ASESIS 197

RESULT 13
 F84169
 hypothetical protein Vng0091c [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84169
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: F84169
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-304 <STO>
 A:Cross-references: GB:AE004437; NID:g10579742; PIDN:AA818722.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0091C

Query Match 96.0%; Score 24; DB 2; Length 304;
 Best Local Similarity 83.3%; Pred. No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 |||||
 DB 229 ASESVS 234

RESULT 14
 T19485
 hypothetical protein C26D10.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19485
 R:Sims, M.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z19129
 A:Accession: T19485
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-349 <WIL>
 A:Cross-references: EMBL:Z54327; PIDN:CAA91121.1; GSPDB:GN00020; CESP:C26D10.3
 A:Experimental source: clone C26D10
 C:Genetics:
 A:Gene: CESP:C26D10.3
 A:Map position: 2
 A:Introns: 47/3; 197/1; 277/3

Query Match 96.0%; Score 24; DB 2; Length 349;
 Best Local Similarity 83.3%; Pred. No. 78;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 |||||
 DB 212 ASESVS 217

RESULT 15
 T13299
 probable helicase - Streptococcus phage phi-O1205
 C:Species: Streptococcus phage phi-O1205
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999

C:Accession: T13299
 R:Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, D.
 Microbiology 143, 3417-3429, 1997
 A:Title: Sequence analysis and characterization of phi O1205, a temperate bacteriophage
 A:Reference number: Z17654; MUID:98048466; PMID:9387220
 A:Accession: T13299
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <STA>
 A:Cross-references: EMBL:U88974; NID:g2444080; PID:g2444090; PIDN:AA079526.1
 A:Experimental source: host Streptococcus thermophilus strain CNR21205

Query Match 96.0%; Score 24; DB 2; Length 443;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 |||||
 DB 225 ASESVS 230

Search completed: October 6, 2004, 16:34:54
 Job time : 9.31579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 75.7895 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-4

Perfect score: 88
Sequence: 1 VWSGGNTDYNTPPT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1980s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|-------------|
| 1 | 88 | 100.0 | 16 | AAW08940 | CDR2 from |
| 2 | 88 | 100.0 | 16 | AAV59312 | Heavy cha |
| 3 | 88 | 100.0 | 16 | AAV59312 | Heavy cha |
| 4 | 88 | 100.0 | 16 | AAV59312 | Heavy cha |
| 5 | 88 | 100.0 | 16 | AAV59312 | Heavy cha |
| 6 | 88 | 100.0 | 119 | AAW08953 | Heavy cha |
| 7 | 88 | 100.0 | 119 | AAW08953 | Heavy cha |
| 8 | 88 | 100.0 | 119 | AAW08953 | Heavy cha |
| 9 | 88 | 100.0 | 119 | AAW08953 | Heavy cha |
| 10 | 88 | 100.0 | 119 | AAW08953 | Heavy cha |
| 11 | 88 | 100.0 | 138 | AAW08947 | Heavy cha |
| 12 | 88 | 100.0 | 138 | AAW08944 | Heavy cha |
| 13 | 88 | 100.0 | 138 | AAW08942 | Heavy cha |
| 14 | 88 | 100.0 | 240 | AAW05133 | Single ch |
| 15 | 88 | 100.0 | 651 | AAW05135 | scFv(225) |
| 16 | 88 | 100.0 | 892 | AAW05140 | scFv2(225) |
| 17 | 88 | 100.0 | 892 | AAW05139 | scFv2(FRP |
| 18 | 88 | 100.0 | 1020 | AAW05141 | scFv2(FRP |
| 19 | 70 | 79.5 | 16 | AAV59312 | Heavy cha |
| 20 | 70 | 79.5 | 114 | AAV59312 | Heavy cha |
| 21 | 65 | 73.9 | 16 | AAU72856 | Anti-NGK2 |
| 22 | 65 | 73.9 | 16 | AAU72856 | Anti-NGK2 |
| 23 | 65 | 73.9 | 117 | ADD25438 | Heavy cha |
| 24 | 65 | 73.9 | 119 | AAE25728 | Humanised |
| 25 | 65 | 73.9 | 119 | AAE25728 | Humanised |

| | | | | | | |
|----|----|------|-----|---|----------|-----------|
| 26 | 65 | 73.9 | 119 | 4 | AAE69675 | Murine mi |
| 27 | 65 | 73.9 | 119 | 4 | AAE69660 | Human lay |
| 28 | 65 | 73.9 | 120 | 5 | AAU72846 | Anti-NGK2 |
| 29 | 65 | 73.9 | 120 | 5 | AAU72854 | Anti-NGK2 |
| 30 | 65 | 73.9 | 127 | 2 | AAO6355 | Peptide c |
| 31 | 65 | 73.9 | 138 | 4 | AAE69656 | Murine mi |
| 32 | 65 | 73.9 | 219 | 2 | AAE69656 | h66-118/h |
| 33 | 65 | 73.9 | 223 | 2 | AAE69656 | h66-111/l |
| 34 | 65 | 73.9 | 228 | 2 | AAE69656 | h66-111/l |
| 35 | 65 | 73.9 | 228 | 2 | AAE69656 | h66-111/l |
| 36 | 65 | 73.9 | 228 | 2 | AAE69656 | h66-111/l |
| 37 | 65 | 73.9 | 235 | 2 | AAE69656 | h66-111/l |
| 38 | 65 | 73.9 | 235 | 2 | AAE69656 | h66-111/l |
| 39 | 65 | 73.9 | 235 | 2 | AAE69656 | h66-111/l |
| 40 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |
| 41 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |
| 42 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |
| 43 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |
| 44 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |
| 45 | 65 | 73.9 | 236 | 2 | AAE69656 | h66-111/l |

ALIGNMENTS

RESULT 1
AAW08940
ID AAW08940 standard; peptide; 16 AA.
XX AC AAW08940;
XX AC AAW08940;

DT 18-SEP-1997 (first entry)

DE CDR2 from heavy chain variable region of reshaped H225 antibody.

XX Complementarity determining region 2; CDR2; heavy chain; variable region;
KW reshaped; human; monoclonal; antibody; H225; epidermal growth factor;
KW EGF; receptor; inhibition; growth; tumour; cell; late stage; prostatic;
KW prostate.

OS Homo sapiens.

XX WO9640210-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009847.

XX PR 07-JUN-1996; 95US-00482982.

XX PR 15-DEC-1995; 95US-00573289.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX (MRCC-) MRC COLLABORATIVE CENT.

XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;

XX WPI; 1997-051897/05.

XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used

XX for inhibiting tumour growth, esp. of late stage prostatic tumour.

XX Claim 1; Page 83; 112pp; English.

XX The present peptide is the complementarity determining region 2 (CDR2)
CC from the heavy chain variable region of the reshaped human monoclonal
CC antibody (MAb) H225, which is specific for the human epidermal growth
CC factor (EGF) receptor. The MAb, or a fragment, can be used to inhibit the
CC growth of tumour cells, especially late stage prostatic tumour cells in
CC humans, optionally conjugated to a cytotoxic agent, especially
CC doxorubicin, taxol or cisplatin, or a signal transduction, ras or cell
CC cycle inhibitor

XX Sequence 16 AA;

Query Match 100.0%; Score 88; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPPT 15
DB 1 VIWSGGNTDYNTPPT 15

RESULT 2

RAY59312
ID AAY59312 standard; peptide; 16 AA.

XX AC AAY59312;

XX XX 07-MAR-2000 (first entry)

XX DE Heavy chain hypervariable region, CDR2.

XX KW Hypervariable region; complementarity determining region; CDR; tumour;
XX KW single chain antibody; growth inhibitor; human; tumorigenesis; therapy;
XX KW protein receptor tyrosine kinase; heavy chain.

XX OS Mus sp.

XX PN WO9960023-A1.

XX PD 25-NOV-1999.

XX PF 14-MAY-1999; 99WO-US010741.

XX PR 15-MAY-1998; 98US-00079612.

XX PR 15-MAY-1998; 98US-0085613P.

XX PR 07-DEC-1998; 98US-00206138.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PA (UABR-) UAB RES FOUND.

XX PI Waksal HW, Saleh MN, Robert F, Buchsbaum DJ;

XX DR WPI; 2000-062440/05.

XX DR N-PSDB; AA248626.

XX PT Treatment of human tumors, using a combination of radiation and a non-
XX PT radiolabeled protein receptor tyrosine kinase inhibitor.

XX PS Disclosure; Page 14; 31pp; English.

XX CC This sequence is the hypervariable region CDR2 (complementarity
XX CC determining region 2) of the heavy chain of a single chain antibody
XX CC derived from the murine antibody 225. The invention relates to a method
XX CC for inhibiting the growth of tumors in human patients by treating with
XX CC an effective amount of a combination of radiation and a non-radiolabelled
XX CC protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
XX CC which can lead to tumorigenesis. The method can be used in the treatment
XX CC of tumors of e.g. breast, lung, colon, kidney, bladder, head and neck,
XX CC ovary, prostate or brain. The administration of a suitable antibody to
XX CC the patient makes the tumour more susceptible to radiotherapy

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 88; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPPT 15
DB 1 VIWSGGNTDYNTPPT 15

RESULT 3

AAB37953

ID AAB37953 standard; protein; 16 AA.

XX AC AAB37953;

XX DT 12-MAR-2001 (first entry)

XX DE Anti-EGFR monoclonal antibody H chain V region CDR2 peptide sequence.
XX KW Refractory tumour growth inhibition; epidermal growth factor receptor;
XX KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
XX KW complementarity determining region; CDR.

XX OS Mus sp.

XX PN WO200069459-A1.

XX PD 23-NOV-2000.

XX PF 01-MAY-2000; 2000WO-US011756.

XX PR 14-MAY-1999; 99US-00312284.

XX PR 13-AUG-1999; 99US-00374028.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Waksal HW;

XX DR WPI; 2001-016160/02.

XX DR N-PSDB; AAC83236.

XX PT Epidermal growth factor receptor/human epidermal growth factor receptor-1
XX PT antagonist for inhibiting the growth of refractory tumors.

XX PS Disclosure; Page 13; 31pp; English.

XX CC This invention relates to a method for inhibiting the growth of
XX CC refractory tumors that are stimulated by a ligand of epidermal growth
XX CC factor receptor (EGFR) in human patients. The method involves treating
XX CC the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
XX CC optionally with a chemotherapeutic agent or radiation. The antagonist can
XX CC be for example a chimeric anti-EGFR monoclonal antibody, C225. The
XX CC EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
XX CC tumours such as tumours of breast, heart, lung, small intestine, colon,
XX CC spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas,
XX CC skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
XX CC liver, preferably squamous cell carcinomas. The present sequence
XX CC represents the heavy chain variable region complementarity determining
XX CC region 2 amino acid sequence of the chimeric anti-EGFR monoclonal
XX CC antibody C225 which is used in an example illustrating the method of the
XX CC invention

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 88; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPPT 15
DB 1 VIWSGGNTDYNTPPT 15

RESULT 4

AAU77785

ID AAU77785 standard; peptide; 16 AA.

XX AC AAU77785;

XX DT 05-JUN-2002 (first entry)

XX DE Mouse heavy chain hypervariable region (CDR2) of 225 antibody.

XX KW Mouse; heavy chain; antibody; hyperproliferative disease;

KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
KW seboretic keratosis; warts; keloid scars; eczema; 255 antibody;
KW hypervariable region; CDR2; EGFR inhibitor.
OS Mus sp.
XX WO200211677-A2.
XX 14-FEB-2002.
XX 09-AUG-2001; 2001WO-US041647.
XX 09-AUG-2000; 2000US-00635974.
XX (IMCL-) IMCLONE SYSTEMS INC.
XX Teufel T;
XX WPI; 2002-257423/30.
XX N-PSDB; ABK11441.
XX Treating a mammal with hyperproliferative disease especially psoriasis,
XX stimulated by ligand of member of epidermal growth factor family of
XX receptors, by administering antagonist of the receptor.
XX Disclosure; Page 12; 28pp; English.
XX This invention relates to a novel method for treating a mammal with
XX hyperproliferative disease stimulated by a ligand of a member of the
XX epidermal growth factor (EGF) family of receptors. The method involves
XX administering an antibody or a defective receptor that is an antagonist
XX of a member of the EGF receptor family, or a combination of the
XX antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
XX The antibody used in the method of the invention acts as an epidermal
XX growth factor receptor (EGFR) antagonist by inhibiting EGFR /HER1
XX phosphorylation. The method of the invention is useful for treating a
XX mammal with hyperproliferative disease such as psoriasis, actinic
XX keratosis, seboretic keratosis, warts, keloid scars and eczema
XX stimulated by a ligand of a member of the EGF family of receptor. This
XX sequence represents the murine anti-EGF 255 antibody heavy chain (VH)
XX hypervariable region (CDR2) used as an inhibitor of EGFR in the method of
XX the invention
XX
XX Sequence 16 AA;
Query Match 100.0%; Score 88; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VIWGGGNTDYNTPFT 15
Db 1 VIWGGGNTDYNTPFT 15
RESULT 5
AAW08939
ID AAW08939 standard; peptide; 18 AA.
XX AAW08939;
XX 18-SEP-1997 (first entry)
XX CDR2 from heavy chain variable region of reshaped H225 antibody.
XX Complementarity determining region 2; CDR2; heavy chain; variable region;
KW reshaped; human; monoclonal; antibody; H225; epidermal growth factor;
KW EGF; receptor; inhibition; growth; tumour; cell; late stage; prostatic;
KW prostate.
XX Homo sapiens.
OS WO9640210-A1.
XX PN
XX

PD 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009847.
XX 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX Claim 1; Page 83; 112pp; English.
XX The present peptide is the complementarity determining region 2 (CDR2)
XX from the heavy chain variable region of the reshaped human monoclonal
XX antibody (MAB) H225, which is specific for the human epidermal growth
XX factor (EGF) receptor. The MAB, or a fragment, can be used to inhibit the
XX growth of tumour cells, especially late stage prostatic tumour cells in
XX humans, optionally conjugated to a cytotoxic agent, especially
XX doxorubicin, taxol or cisplatin, or a signal transduction, ras or cell
XX cycle inhibitor
XX Sequence 18 AA;
Query Match 100.0%; Score 88; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VIWGGGNTDYNTPFT 15
Db 2 VIWGGGNTDYNTPFT 15
RESULT 6
AAW08953
ID AAW08953 standard; protein; 119 AA.
XX AC AAW08953;
XX 18-SEP-1997 (first entry)
XX Heavy chain variable region of 225RD antibody.
XX Heavy chain; reshaped; monoclonal; antibody; 225RD; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX WO9640210-A1.
XX PN

us-09-635-974a-4.rag

Thu Oct 14 09:36:56 2004

```
XX 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009847.
XX 07-JUN-1995; 95US-00482982.
XX 15-DEC-1995; 95US-00573289.
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX Claim 30; Fig 22; 112pp; English.
XX The present sequence is the heavy chain variable region of the reshaped
XX human monoclonal antibody (MAB) H225, 225RHD. The MAB is specific for the
XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX be used to inhibit the growth of tumour cells, especially late stage
XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX agent, especially doxorubicin, taxol or cisplatin, or a signal
XX transduction, ras or cell cycle inhibitor
XX Sequence 119 AA;
XX Query Match 100.0%; Score 88; DB 2; Length 119;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-06;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 VIWSGGNTDYNTPFT 15
XX DB 50 VIWSGGNTDYNTPFT 64
XX
XX RESULT 7
XX AAW08950
XX ID AAW08950 standard; protein; 119 AA.
XX AC AAW08950;
XX DT 18-SEP-1997 (first entry)
XX DE Heavy chain variable region of 225RA antibody.
XX KW Heavy chain; reshaped; monoclonal; antibody; 225RA; human;
XX KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX KW late stage; prostatic; prostate; variable region; framework;
XX KW complementarity determining region; CDR.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..30
XX FT /label= framework_1
XX FT Region 31..35
XX FT /label= CDR_1
XX FT Region 36..49
XX FT /label= framework_2
XX FT Region 50..65
XX FT /label= CDR_2
XX FT Region 66..97
XX FT /label= framework_3
XX FT Region 98..108
XX FT /label= CDR_3
XX FT Region 109..119
XX FT /label= framework_4
XX PN WO9640210-A1.
```

```
XX 19-DEC-1996.
XX 07-JUN-1996; 96WO-US009847.
XX 07-JUN-1995; 95US-00482982.
XX 15-DEC-1995; 95US-00573289.
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX Claim 30; Fig 22; 112pp; English.
XX The present sequence is the heavy chain variable region of the reshaped
XX human monoclonal antibody (MAB) H225, 225RHA. The MAB is specific for the
XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX be used to inhibit the growth of tumour cells, especially late stage
XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX agent, especially doxorubicin, taxol or cisplatin, or a signal
XX transduction, ras or cell cycle inhibitor
XX Sequence 119 AA;
XX Query Match 100.0%; Score 88; DB 2; Length 119;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-06;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 VIWSGGNTDYNTPFT 15
XX DB 50 VIWSGGNTDYNTPFT 64
XX
XX RESULT 8
XX AAW08952
XX ID AAW08952 standard; protein; 119 AA.
XX AC AAW08952;
XX DT 18-SEP-1997 (first entry)
XX DE Heavy chain variable region of 225RC antibody.
XX KW Heavy chain; reshaped; monoclonal; antibody; 225RC; human;
XX KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX KW late stage; prostatic; prostate; variable region; framework;
XX KW complementarity determining region; CDR.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..30
XX FT /label= framework_1
XX FT Region 31..35
XX FT /label= CDR_1
XX FT Region 36..49
XX FT /label= framework_2
XX FT Region 50..65
XX FT /label= CDR_2
XX FT Region 66..97
XX FT /label= framework_3
XX FT Region 98..108
XX FT /label= CDR_3
XX FT Region 109..119
XX FT /label= framework_4
XX PN WO9640210-A1.
```

```

XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009847.
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (MRCC-) MRC COLLABORATIVE CENT.
XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX DR WPI; 1997-051897/05.
XX PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX PS Claim 30; Fig 22; 112pp; English.
XX CC The present sequence is the heavy chain variable region of the reshaped
XX CC human monoclonal antibody (MAB) H225, 225RHC. The MAB is specific for the
XX CC human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX CC be used to inhibit the growth of tumour cells, especially late stage
XX CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX CC agent, especially doxorubicin, taxol or cisplatin, or a signal
XX CC transduction, ras or cell cycle inhibitor
XX CC
XX SQ Sequence 119 AA;

Query Match 100.0%; Score 88; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPPT 15
Db 50 VWSGGNTDYNTPPT 64

RESULT 9
AAW08951
ID AAW08951 standard; protein; 119 AA.
AC AAW08951;
XX 18-SEP-1997 (first entry)
XX Heavy chain variable region of 225RB antibody.
XX Heavy chain; reshaped; monoclonal; antibody; 225RB; human;
XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX late stage; prostatic; prostate; variable region; framework;
XX complementarity determining region; CDR.
XX Homo sapiens.
XX Key Location/Qualifiers
XX FH 1..30
XX FT /label= framework_1
XX FT 31..35
XX FT /label= CDR_1
XX FT 36..49
XX FT /label= framework_2
XX FT 50..65
XX FT /label= CDR_2
XX FT 66..97
XX FT /label= framework_3
XX FT 98..108
XX FT /label= CDR_3
XX FT 109..119
XX FT /label= framework_4
XX PN WO9640210-A1.

Query Match 100.0%; Score 88; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPPT 15
Db 50 VWSGGNTDYNTPPT 64

RESULT 10
AAW08954
ID AAW08954 standard; protein; 119 AA.
AC AAW08954;
XX 18-SEP-1997 (first entry)
XX Heavy chain variable region of 225RE antibody.
XX Heavy chain; reshaped; monoclonal; antibody; 225RE; human;
XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX late stage; prostatic; prostate; variable region; framework;
XX complementarity determining region; CDR.
XX Homo sapiens.
XX Key Location/Qualifiers
XX FH 1..30
XX FT /label= framework_1
XX FT 31..35
XX FT /label= CDR_1
XX FT 36..49
XX FT /label= framework_2
XX FT 50..65
XX FT /label= CDR_2
XX FT 66..97
XX FT /label= framework_3
XX FT 98..108
XX FT /label= CDR_3
XX FT 109..119
XX FT /label= framework_4
XX PN WO9640210-A1.

```

XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 XX 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 XX (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Claim 30; Fig 22; 112pp; English.
 XX The present sequence is the heavy chain variable region of the reshaped
 XX human monoclonal antibody (MAB) H225, 225RHA. The MAB is specific for the
 XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
 XX be used to inhibit the growth of tumour cells, especially late stage
 XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 XX agent, especially doxorubicin, taxol or cisplatin, or a signal
 XX transduction, ras or cell cycle inhibitor
 XX Sequence 119 AA;
 XX Query Match 100.0%; Score 88; DB 2; Length 119;
 XX Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VWSGGTNDYNTPT 15
 DB 50 VWSGGTNDYNTPT 64
 RESULT 11
 AA08947
 ID AAW08947 standard; protein; 138 AA.
 AC AAW08947;
 XX 18-SEP-1997 (first entry)
 DE Heavy chain variable region of 225RA antibody.
 XX Heavy chain; reshaped; monoclonal; antibody; 225RA; human;
 XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 XX late stage; prostatic; prostate; variable region.
 XX Homo sapiens.
 XX WO9640210-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 XX 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 XX (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 XX N-PSDB; AAT49346.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used

PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Claim 31; Fig 20; 112pp; English.
 XX The present sequence is the heavy chain variable region of the reshaped
 XX human monoclonal antibody (MAB) H225, 225RHA. The MAB is specific for the
 XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
 XX be used to inhibit the growth of tumour cells, especially late stage
 XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 XX agent, especially doxorubicin, taxol or cisplatin, or a signal
 XX transduction, ras or cell cycle inhibitor
 XX Sequence 138 AA;
 XX Query Match 100.0%; Score 88; DB 2; Length 138;
 XX Best Local Similarity 100.0%; Pred. No. 3.2e-06;
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VWSGGTNDYNTPT 15
 DB 69 VWSGGTNDYNTPT 83
 RESULT 12
 AA08944
 ID AAW08944 standard; protein; 138 AA.
 AC AAW08944;
 XX 18-SEP-1997 (first entry)
 DE Heavy chain variable region of C225 antibody.
 XX Heavy chain; murine; mouse; monoclonal; antibody; C225; human;
 XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 XX late stage; prostatic; prostate; variable region; chimeric.
 XX Mus; spp.
 XX Homo sapiens.
 XX Synthetic.
 XX Chimeric.
 XX WO9640210-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 XX 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 XX (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 XX N-PSDB; AAT49343.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Example IV; Fig 16; 112pp; English.
 XX The present sequence is the heavy chain variable region of the chimeric
 XX monoclonal antibody (MAB) C225, which is specific for the human epidermal
 XX growth factor (EGF) receptor. The MAB, or a fragment, can be used to
 XX inhibit the growth of tumour cells, especially late stage prostatic
 XX tumour cells in humans, optionally conjugated to a cytotoxic agent,
 XX especially doxorubicin, taxol or cisplatin, or a signal transduction, ras
 XX or cell cycle inhibitor
 XX Sequence 138 AA;
 XX SQ

Query Match 100.0%; Score 88; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPPT 15
Db 69 VIWGGNTDYNTPPT 83

RESULT 13
AAW08942
ID AAW08942 standard; protein; 138 AA.
AC AAW08942;
XX
XX 27-AUG-2003 (revised)
DT 18-SEP-1997 (first entry)
XX
XX Heavy chain variable region of M225 antibody.
XX
XX Heavy chain; murine; mouse; monoclonal; antibody; M225; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XW late stage; prostatic; prostate; variable region.
XX
XX Mus sp.
OS
XX
XX WO9640210-A1.
PN
PD
PF 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009847.
PF
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NT, Giorgio NA, Jones ST, Saldanha JW;
PI
XX
XX WPI; 1997-051897/05.
DR N-PSDB; AAT49340.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Example IV; Fig 14; 112pp; English.
PS
XX The present sequence is the heavy chain variable region of the murine
CC monoclonal antibody (Mab) M225, which is specific for the human epidermal
CC growth factor (EGF) receptor. The Mab, or a fragment, can be used to
CC inhibit the growth of tumour cells, especially late stage prostatic
CC tumour cells in humans, optionally conjugated to a cytotoxic agent,
CC especially doxorubicin, taxol or cisplatin, or a signal transduction, ras
CC or cell cycle inhibitor. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 138 AA;
SQ

Query Match 100.0%; Score 88; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPPT 15
Db 69 VIWGGNTDYNTPPT 83

RESULT 14
AAW05133
ID AAW05133 standard; protein; 240 AA.
XX
XX AAW05133;

XX 29-JAN-1997 (first entry)
DT
XX Single chain antibody scFv(225).
DE
XX
XX Single chain antibody; scFv; monoclonal antibody; MAB; EGF;
KW epidermal growth factor; receptor; antitumour; cancer; therapy.
XX
XX Mus; sp.
OS Synthetic.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH Region 1..119
FT /label= VH region
FT /note= "monoclonal antibody 225 VH"
FT Peptide 120..133
FT /label= Linker
FT /note= "synthetic spacer peptide"
FT 134..240
FT /label= VL
FT /note= "monoclonal antibody 225 VL"
XX
XX EP739984-A1.
PN
XX
XX 30-OCT-1996.
PD
XX
XX 26-APR-1995; 95EP-00106275.
PF
XX
XX 26-APR-1995; 95EP-00106275.
PR
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
PA
XX
XX Wels W, Schmidt M, Groner B;
PI
XX
XX WPI; 1996-478748/48.
DR N-PSDB; AAT42033.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
XX Example 2; Page 17; 52pp; English.
PS
XX
XX scFv(225) (AAW05133) comprises the single-chain binding region of murine
CC monoclonal antibody 225, which is specific for the human epidermal growth
CC factor receptor. It is encoded by plasmid pMW152-225 (see also AAT42033),
CC constructed by cloning Mab 225 VH and VL region cDNAs into plasmid
CC pMW152. Novel bivalent proteins (see also AAW05134-44), some of them
CC including scFv(225) and an effector e.g. cytotoxin, can be produced in
CC bacterial host cells, and are useful as antitumour agents
XX
XX Sequence 240 AA;
SQ

Query Match 100.0%; Score 88; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPPT 15
Db 50 VIWGGNTDYNTPPT 64

RESULT 15
AAW05135
ID AAW05135 standard; protein; 651 AA.
XX
XX AAW05135;
AC
XX
XX 29-JAN-1997 (first entry)
DT
XX
XX scFv(225)-ETA fusion protein.
DE
XX

us-09-635-974a-4.rag

Thu Oct 14 09:36:56 2004

KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF;
 KW epidermal growth factor; receptor; plasmid pSW202-225; cancer; therapy;
 KW antitumour; exotoxin A; ETA.

XX Mus sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers
 FH 1..21
 FT Peptide /label= Sig_peptide
 FT /note= "ompA signal peptide"
 FT 22..38
 FT Peptide /label= Spacer
 FT 39..278
 FT Protein /label= scFv(225)
 FT 279..289
 FT Peptide /label= Spacer
 FT 290..651
 FT Protein /label= ETA
 FT /note= "exotoxin A amino acids 252-613"

XX EP739984-A1.

PN 30-OCT-1996.

XX 26-APR-1995; 95EP-00106275.

XX 26-APR-1995; 95EP-00106275.

XX (SANT-) SAN TUMORFORSCHUNGS GMBH.

XX Wels W, Schmidt M, Groner B;

XX WPI; 1996-478748/48.

XX N-PSDB; AAT42035.

XX Bivalent fusion proteins that bind epidermal growth factor receptor or
 PT analogues - and comprise at least two different cell surface binding
 PT domain(s), useful for tumour therapy.

XX Example 7; Page 19-20; 52pp; English.

XX scFv(225)-ETA (AAW05135) comprises the single-chain binding region (see
 CC also AAW05133) of murine monoclonal antibody 225, which is specific for
 CC human epidermal growth factor receptor, joined to exotoxin A (ETA). It is
 CC encoded by plasmid pSW202-225 (see also AAT42035) obtcd. by ligating an
 CC scFv(225) gene (AAT42033) into plasmid pSW200 contg. the Pseudomonas
 CC aeruginosa PAK ETA gene. The construct can be used to produce novel
 CC bivalent fusion proteins (see also AAW05136-44) in bacterial host cells,
 CC for use as antitumour agents

XX Sequence 651 AA;

Query Match 100.0%; Score 88; DB 2; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15

DB 88 VIWSGGNTDYNTPFT 102

Search completed: October 6, 2004, 16:29:43
 Job time : 76.7895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: October 6, 2004, 16:23:59 ; Search time 15.7895 Seconds
(without alignments)
91.382 Million cell updates/sec

Title: US-09-635-974A-4

Perfect score: 88

Sequence: 1 VWSGGNTDYNTPFT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:**

1: PIR1:**

2: PIR2:**

3: PIR3:**

4: PIR4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 67 | 76.1 | 109 | 2 PH1026 | Ig heavy chain V r |
| 2 | 65 | 73.9 | 90 | 2 A49042 | Ig heavy chain V r |
| 3 | 65 | 73.9 | 100 | 2 A25913 | Ig heavy chain pre |
| 4 | 65 | 73.9 | 115 | 2 S26470 | Ig heavy chain V r |
| 5 | 65 | 73.9 | 115 | 2 S11107 | Ig heavy chain V r |
| 6 | 65 | 73.9 | 116 | 1 G1MS10 | Ig heavy chain pre |
| 7 | 65 | 73.9 | 116 | 2 A33932 | Ig mu chain precu |
| 8 | 65 | 73.9 | 116 | 2 PL0198 | anti-DNA autoantib |
| 9 | 65 | 73.9 | 121 | 2 D30560 | Ig heavy chain V r |
| 10 | 65 | 73.9 | 127 | 2 B31807 | Ig heavy chain V r |
| 11 | 65 | 73.9 | 135 | 2 S31913 | Ig gamma-2A chain |
| 12 | 65 | 73.9 | 140 | 2 S14238 | Ig gamma-1 chain p |
| 13 | 64 | 72.7 | 114 | 2 S26321 | Ig heavy chain V r |
| 14 | 60 | 68.2 | 107 | 2 S14506 | Ig heavy chain V r |
| 15 | 60 | 68.2 | 121 | 2 S33131 | Ig heavy chain V r |
| 16 | 60 | 68.2 | 140 | 2 S55028 | Ig heavy chain V r |
| 17 | 59 | 67.0 | 117 | 2 S38563 | Ig heavy chain V r |
| 18 | 58 | 65.9 | 139 | 2 A32456 | Ig heavy chain pre |
| 19 | 54 | 61.4 | 337 | 2 S26950 | NADH2 dehydrogenas |
| 20 | 54 | 61.4 | 653 | 2 T14241 | NADH2 dehydrogenas |
| 21 | 53 | 60.2 | 95 | 2 S17605 | Ig heavy chain V r |
| 22 | 53 | 60.2 | 100 | 2 S14490 | Ig heavy chain V r |
| 23 | 53 | 60.2 | 106 | 2 S26322 | Ig heavy chain V r |
| 24 | 53 | 60.2 | 106 | 2 S14489 | Ig heavy chain V r |
| 25 | 53 | 60.2 | 107 | 2 S14492 | Ig heavy chain V r |
| 26 | 53 | 60.2 | 107 | 2 S14493 | Ig heavy chain V r |
| 27 | 53 | 60.2 | 109 | 2 PH1025 | Ig heavy chain V r |
| 28 | 53 | 60.2 | 109 | 2 S11109 | Ig heavy chain V r |
| 29 | 53 | 60.2 | 112 | 2 S11098 | Ig heavy chain V r |

30 53 60.2 112 2 S11100 Ig heavy chain V r
31 53 60.2 112 2 S11108 Ig heavy chain V r
32 53 60.2 113 2 S11101 Ig heavy chain V r
33 53 60.2 114 2 S11099 Ig heavy chain V r
34 53 60.2 114 2 S11104 Ig heavy chain V r
35 53 60.2 114 2 S11106 Ig heavy chain V r
36 53 60.2 114 2 S11105 Ig heavy chain V r
37 53 60.2 115 2 S11103 Ig heavy chain V r
38 53 60.2 116 2 S11102 Ig heavy chain V r
39 53 60.2 117 2 S10111 Ig heavy chain V r
40 53 60.2 120 2 PL0087 Ig heavy chain V r
41 53 60.2 122 2 A49049 Ig heavy chain V r
42 53 60.2 141 2 S52446 Ig heavy chain V r
43 53 60.2 301 2 H75493 arginase - Deinoco
44 52 59.1 107 2 S14491
45 52 59.1 110 2 PH1024 Ig heavy chain V r

ALIGNMENTS

RESULT 1

PH1026

Ig heavy chain V region (clone 163-c3) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PH1026

R:Fillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH1026

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-109 <TIL>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 76.1%; Score 67; DB 2; Length 109;

Best Local Similarity 78.6%; Pred. No. 0.0018;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPF 14

DB 50 MWSGGNTDYNAAF 63

RESULT 2

A49042

Ig heavy chain V region (anti-DNA) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999

C:Accession: A49042

R:Taki, S.; Hirose, S.; Kinoshita, K.; Nishimura, H.; Shimamura, T.; Hamuro, J.; Shirai,

Eur. J. Immunol. 22, 987-992, 1992

A:Title: Sonatically mutated IGG anti-DNA antibody clonally related to germ-line encoded

A:Reference number: A49042; MUID:92201320; PMID:1551410

A:Accession: A49042

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-90 <TAK>

A:Cross-references: GB:892270; NID:g247954; PIDN:AAB21904.1; PID:g247955

A:Experimental source: NZB, NZW, liver

A>Note: sequence extracted from NCBI backbone (NCBIN:92270, NCBI:92271)

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:8-90/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 2; Length 90;

Best Local Similarity 78.6%; Pred. No. 0.003;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 43 VIWSGGSTDYNAAF 56

RESULT 3

A35913
 Ig heavy chain precursor V region (BFL2) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 21-Jan-2000
 C:Accession: A25913
 R:Lawler, A.M.; Lin, P.S.; Gearhart, P.J.
 Proc. Natl. Acad. Sci. U.S.A. 84, 2454-2458, 1987
 A:Title: Adult B-cell repertoire is biased toward two heavy-chain variable-region genes
 A:Reference number: A94148; MUID:87175692; PMID:3104915
 A:Accession: A25913
 A:Molecule type: DNA
 A:Residues: 1-100 <LAW>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:19-100/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 2; Length 100;
 Best Local Similarity 78.6%; Pred. No. 0.0033; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 54 VIWSGGSTDYNAAF 67

RESULT 4

S26470
 Ig heavy chain V region - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
 C:Accession: S26470
 R:Kavaler, J.
 submitted to the EMBL Data Library, April 1991
 A:Reference number: S26459
 A:Accession: S26470
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-115 <KAV>
 A:Cross-references: EMBL:X59109; NID:G51948; PIDN:CAA41835.1; PID:G51949
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 2; Length 115;
 Best Local Similarity 78.6%; Pred. No. 0.0038; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 50 VIWSGGSTDYNAAF 63

RESULT 5

S11107
 Ig heavy chain V region (clone NQ2-45.10.4) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jan-2000
 C:Accession: S11107
 R:Kaattinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
 Nature 304, 320-324, 1983
 A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenylloxazolone
 A:Reference number: S07331; MUID:83271467; PMID:6877353
 A:Accession: S11107
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-115 <KAA>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 2; Length 115;
 Best Local Similarity 78.6%; Pred. No. 0.0038; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 50 VIWSGGSTDYNAAF 63

RESULT 6

G1M510
 Ig heavy chain precursor V region (MC101) - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 22-Jun-1999
 C:Accession: A02096
 R:Kataoka T.; Nikaido, T.; Miyata, T.; Moriaki, K.; Honjo, T.
 J. Biol. Chem. 257, 277-285, 1982
 A:Title: The nucleotide sequences of rearranged and germ-line immunoglobulin V-H genes of
 A:Reference number: A02096; MUID:82075900; PMID:6273429
 A:Accession: A02096
 A:Molecule type: DNA
 A:Residues: 1-116 <XAT>
 A:Cross-references: GB:J00502; NID:G196070; PIDN:AAA38515.1; PID:G196071
 A:Note: the sequence was determined from the germline gene
 C:Genetics:

A:Introns: 16/1

C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-116/Product: Ig heavy chain V region (MC101) #status predicted <MAT>
 F:34-116/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 1; Length 116;
 Best Local Similarity 78.6%; Pred. No. 0.0039;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 69 VIWSGGSTDYNAAF 82

RESULT 7

A33932
 Ig mu chain precursor V region (D23) - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 09-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 23-May-1997
 C:Accession: A33932
 R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.
 Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
 A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated germ-line
 A:Reference number: A33932; MUID:89282823; PMID:2499887
 A:Accession: A33932
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-116 <BAC>

A:Cross-references: GB:M27107
 A:Note: the authors translated the codon CTA for residue 18 as Thr
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:34-116/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 65; DB 2; Length 116;
 Best Local Similarity 78.6%; Pred. No. 0.0039;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||:|||||
 Db 69 VIWSGGSTDYNAAF 82

RESULT 8

PLO198
 anti-DNA autoantibody BV16-13, heavy chain V region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
 C:Accession: PLO198
 R:Smith, R.G.; Voss Jr., E.W.
 Mol. Immunol. 27, 463-470, 1990
 A:Title: Variable region primary structures of monoclonal anti-DNA autoantibodies from N
 A:Reference number: PLO198; MUID:90309768; PMID:2114528
 A:Accession: PLO198
 A:Molecule type: mRNA
 A:Residues: 1-116 <SMI>
 A:Cross-references: GB:X53638; NID:950191; PIDN:CAA37689.1; PID:930137
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 F:15-97/Domain: immunoglobulin homology <IMM>
 F:31-35/Region: complementarity-determining 1
 F:50-65/Region: complementarity-determining 2
 F:95-105/Region: complementarity-determining 3
 F:95-101/Region: D region
 F:102-113/Region: JH region
 Query Match 73.9%; Score 65; DB 2; Length 116;
 Best Local Similarity 78.6%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPF 14
 |||||
 DB 50 VIWSDGRIDYNVFP 53
 |||||
 RESULT 9
 D30560
 IG heavy chain V region (36.1.2D) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 23-Mar-1989 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999
 C:Accession: D30560
 R:Matsuda, T.; Kabat, E.A.
 J. Immunol. 142, 863-870, 1989
 A:Title: Variable region cDNA sequences and antigen binding specificity of mouse monoclo
 A:Reference number: A30560; MUID:89110062; PMID:2464028
 A:Accession: D30560
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-121 <MAT>
 A:Cross-references: GB:M24271; NID:G195617; PIDN:AAA38372.1; PID:G195618
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>
 Query Match 73.9%; Score 65; DB 2; Length 121;
 Best Local Similarity 78.6%; Pred. No. 0.0041;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPF 14
 |||||
 DB 50 VIWGGNTDYNAAP 63
 |||||
 RESULT 10
 B31807
 IG heavy chain V region (PAC1) - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 20-Jul-1989 #sequence_revision 20-Jul-1989 #text_change 23-May-1997
 C:Accession: B31807
 R:Taub, R.; Gould, R.J.; Garsky, V.M.; Ciccarone, T.M.; Hoxie, J.; Friedman, P.A.; Shatt
 J. Biol. Chem. 264, 259-265, 1989
 A:Title: A monoclonal antibody against the platelet fibrinogen receptor contains a sequ
 A:Reference number: A31807; MUID:89079661; PMID:2909518
 A:Accession: B31807
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-127 <TAU>

C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>
 Query Match 73.9%; Score 65; DB 2; Length 127;
 Best Local Similarity 78.6%; Pred. No. 0.0043;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPF 14
 |||||
 DB 50 VIWGGSTDYNAAP 63
 |||||
 RESULT 11
 S31913
 IG gamma-2A chain precursor - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 23-Jul-1999
 C:Accession: S31913
 R:Bespalov, I.A.; Hiyonov, P.A.; Lukashevich, L.V.; Lunev, V.E.; Tribush, S.S.; Gaponov
 submitted to the EMBL Data Library, January 1993
 A:Reference number: S31913
 A:Accession: S31913
 A:Molecule type: mRNA
 A:Residues: 1-135 <BES>
 A:Cross-references: EMBL:X70822; NID:957921; PIDN:CAA50153.1; PID:G57922
 A:Experimental source: strain BALB/c
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:1-11/Domain: signal sequence #status predicted <SIG>
 F:12-135/Product: IG gamma-2A chain (fragment) #status predicted <MAT>
 F:34-116/Domain: immunoglobulin homology <IMM>
 Query Match 73.9%; Score 65; DB 2; Length 135;
 Best Local Similarity 78.6%; Pred. No. 0.0046;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPF 14
 |||||
 DB 69 VIWGGSTDYNAAP 82
 |||||
 RESULT 12
 S14238
 IG gamma-1 chain precursor (15C5) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S14238
 R:Vandamme, A.M.; Bulens, F.; Bernar, H.; Nelles, L.; Lijnen, R.H.; Collen, D.
 Eur. J. Biochem. 192, 767-775, 1990
 A:Title: Construction and characterization of a recombinant murine monoclonal antibody c
 A:Reference number: S14236; MUID:91006173; PMID:2209622
 A:Accession: S14238
 A:Molecule type: mRNA
 A:Residues: 1-140 <VAN>
 A:Cross-references: EMBL:X56392; NID:951619; PIDN:CAA39803.1; PID:G747853
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:26-108/Domain: immunoglobulin homology <IMM>
 Query Match 73.9%; Score 65; DB 2; Length 140;
 Best Local Similarity 78.6%; Pred. No. 0.0047;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPF 14
 |||||
 DB 61 VIWGGSTDYNAAP 74
 |||||
 RESULT 13
 S26321
 IG heavy chain V region - mouse (fragment)
 C:Species: Mus musculus (house mouse)

C>Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 21-Jul-2000
 C:Accession: S26321
 R:Stark, S.E.; Caton, A.J.
 J:Exp. Med. 174: 633-624, 1991
 A:Title: Antibodies that are specific for a single amino acid interchange in a protein
 A:Reference number: S26309; MUID:91341421; PMID:1908510
 A:Accession: S26321
 A:Molecule type: mRNA
 A:Residues: 1-114 <STA>
 A:Cross-references: EMBL:X59176; NID:G52051; PIDN:CAA41886.1; PID:G4379205
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:11-93/Domain: immunoglobulin homology <IMM>

Query Match 72.7%; Score 64; DB 2; Length 114;
 Best Local Similarity 78.6%; Pred. No. 0.0055;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||
 DB 46 VIWSGGTDDYNAAF 59

RESULT 14

S14506
 Ig heavy chain V region - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 C:Accession: S14506
 R:Chen, Q.; Stenzel-Poore, M.; Rittenberg, M.B.
 submitted to the EMBL Data Library, March 1991
 A:Description: Natural polyclonal antibodies differ from Ag-induced antibodies in VH C
 A:Reference number: S14484
 A:Accession: S14506
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-107 <CHE>
 A:Cross-references: EMBL:X58651; NID:G51291; PIDN:CAA41508.1; PID:G51292
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 68.2%; Score 60; DB 2; Length 107;
 Best Local Similarity 71.4%; Pred. No. 0.022;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||
 DB 50 VIWSGGTDDYNAAF 63

RESULT 15

S3131
 Ig heavy chain V region - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
 C:Accession: S3131
 R:Tempest, P.R.; Barbanti, E.; Bremner, P.; Carr, F.J.; Ghislieri, M.; Rifaldi, B.; Marc
 submitted to the EMBL Data Library, May 1993
 A:Description: A humanized anti-tumor necrosis factor alpha monoclonal antibody that act
 A:Reference number: S3131
 A:Accession: S3131
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-121 <TEM>
 A:Cross-references: EMBL:Z22669; NID:G297471; PIDN:CAA90378.1; PID:G297472
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 68.2%; Score 60; DB 2; Length 121;
 Best Local Similarity 71.4%; Pred. No. 0.025;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
 |||||
 DB 50 VIWSGGTDDYNAAF 63
 Search completed: October 6, 2004, 16:34:46
 Job time : 16.7895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 9.21053 seconds
(without alignments)
84.800 Million cell updates/sec

Title: US-09-635-974A-4

Perfect score: 88

Sequence: 1 VIWSGGNTDNTPT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 65 | 73.9 | 116 | 1 HV45 MOUSE | P01821 mus musculus |
| 2 | 54 | 61.4 | 653 | 1 NUSM_TIRU | O01561 trichophyto |
| 3 | 51 | 58.0 | 852 | 1 NUSM_PODAN | P20679 podospora a |
| 4 | 51 | 58.0 | 657 | 1 NUSM_EMENI | P11628 emericeella |
| 5 | 50.5 | 57.4 | 507 | 1 YG46 YEAST | P53301 saccharomyc |
| 6 | 50 | 56.8 | 322 | 1 ARG1 HUMAN | P05089 homo sapien |
| 7 | 50 | 56.8 | 323 | 1 ARG1 MOUSE | O61176 mus musculus |
| 8 | 50 | 56.8 | 323 | 1 ARG1 RAT | P07824 rattus norv |
| 9 | 47 | 53.4 | 115 | 1 HV45 MOUSE | P01820 mus musculus |
| 10 | 47 | 53.4 | 144 | 1 HV43 MOUSE | P01819 mus musculus |
| 11 | 46 | 52.3 | 715 | 1 NUSM_NEUCR | P05510 neurospora |
| 12 | 45 | 51.1 | 168 | 1 PIG1 DROME | P26023 drosophila |
| 13 | 45 | 51.1 | 354 | 1 ARG2 MOUSE | O08691 mus musculus |
| 14 | 45 | 51.1 | 354 | 1 ARG2 RAT | O08701 rattus norv |
| 15 | 45 | 51.1 | 360 | 1 ARG1_XENLA | O91553 xenopus lae |
| 16 | 45 | 51.1 | 360 | 1 ARG2_XENLA | O91554 xenopus lae |
| 17 | 45 | 51.1 | 323 | 1 ARG1_XENLA | O91555 xenopus lae |
| 18 | 44 | 50.0 | 323 | 1 ARG1_RANCA | P49900 rana catesb |
| 19 | 44 | 50.0 | 354 | 1 ARG2 HUMAN | P78540 homo sapien |
| 20 | 42.5 | 48.3 | 704 | 1 CDGT_BACOH | P27036 bacillus oh |
| 21 | 42 | 47.7 | 241 | 1 H1S4_BIFLO | O8455 bifidobacte |
| 22 | 42 | 47.7 | 316 | 1 ARG1_XENLA | P30759 xenopus lae |
| 23 | 42 | 47.7 | 358 | 1 ARG1_NEUCR | P33280 neurospora |
| 24 | 42 | 47.7 | 467 | 1 GAC2_HUMAN | P18507 homo sapien |
| 25 | 42 | 47.7 | 475 | 1 GAC2_BOVIN | P22300 bos taurus |
| 26 | 42 | 47.7 | 654 | 1 NUSM_RHIST | P50367 rhizopus st |
| 27 | 41 | 46.6 | 639 | 1 NPT2 SHEEP | O97704 ovis aries |
| 28 | 41 | 46.6 | 796 | 1 VP35_HUMAN | O96qk1 homo sapien |
| 29 | 41 | 46.6 | 796 | 1 VP35_MOUSE | O96qk1 mus musculus |
| 30 | 40 | 45.5 | 283 | 1 LECA_SARPE | P05047 sarcophaga |
| 31 | 40 | 45.5 | 440 | 1 DCO_DROME | O76324 drosophila |
| 32 | 40 | 45.5 | 1077 | 1 AT10_HUMAN | O9H324 homo sapien |
| 33 | 40 | 45.5 | 1456 | 1 MANR_HUMAN | P22897 homo sapien |

```

34 39.5 44.9 359 1 LPFD_SALTY
35 39.5 44.9 703 1 CDGT_BACS2
36 39.5 44.9 1240 1 YNU1_YEAST
37 39 44.3 168 1 COAD_WIGER
38 39 44.3 258 1 TT2_ARATH
39 39 44.3 259 1 Y895_NEIMB
40 39 44.3 259 1 YB14_NEIMA
41 39 44.3 260 1 MYC1_CHUNI
42 39 44.3 355 1 AROB_STRPN
43 39 44.3 355 1 AROB_STRR6
44 39 44.3 364 1 RIBB_PHOLE
45 39 44.3 576 1 IL1R_MOUSE

P43663 salmonella
P31746 bacillus sp
P53935 saccharomyc
Q82r5 wigleswort
Q9fja2 arabidopsis
Q9fzu5 neisseria m
Q9juv9 neisseria m
Q01511 chlorella v
Q97q56 streptococc
Q8dpd1 streptococc
Q02008 photobacter
P13504 mus musculus

```

ALIGNMENTS

```

RESULT 1
HV45_MOUSE STANDARD; PRT; 116 AA.
ID HV45_MOUSE
AC P01821
DT 21-JUL-1986 (Rel. 01, Created)
DE 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE IG heavy chain V region MC101 precursor.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82075900; PubMed=6273429;
RA Kataoka T., Nikaido T., Miyata T., Moriaki K., Honjo T.;
RT "The nucleotide sequences of rearranged and germline immunoglobulin
RT VH genes of a mouse myeloma MC101 and evolution of VH genes in
RT mouse.";
RL J. Biol. Chem. 257:277-285(1982).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00502; AAA38515.1; -.
CC PIR; A02096; GIMSL0.
CC HSSP; P01772; 2FB4.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003596; IG_v.
CC Pfam; PF00047; ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS5083; IG LIKE; 1.
CC Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 19
FT DOMAIN 20 >116 IG HEAVY CHAIN V REGION MC101.
FT NON_TER 116 116 IG-LIKE.
SQ SEQUENCE 116 AA; 12593 MW; 8079A6EB7C552B3E CRC64;

Query Match 73.9%; Score 65; DB 1; Length 116;
Best Local Similarity 78.6%; Pred. No. 0.0022;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDNTPT 14
DQ 69 VIWSGGSTDNTAA 82

RESULT 2
NUSM_TIRU

```

```

ID NUSM TRIRU STANDARD; PRT; 653 AA.
AC Q01561; Q979N1;
DT 01-JUN-1994 (Rel. 29, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NADH5.
OS Trichophyton rubrum.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Arthrodermataceae; mitosporic Arthrodermataceae;
OC Trichophyton.
OX NCBI_TaxID=5551;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 1817.89;
RX MEDLINE=99146870; PubMed=10022946;
RA de Bievre C., Dujon B.;
RT "Organisation of the mitochondrial genome of Trichophyton rubrum III.
RT DNA sequence analysis of the NADH dehydrogenase subunits 1, 2, 3, 4,
RT 5 and the cytochrome b gene.";
RL Curr. Genet. 35:30-35(1999).
RN [2]
RP SEQUENCE OF 1-337 FROM N.A.
RC STRAIN=IP 1817.89;
RX MEDLINE=92405249; PubMed=1126416;
RA de Bievre C., Dujon B.;
RT "Mitochondrial DNA sequence analysis of the cytochrome oxidase
RT subunit I and II genes, the ATPase9 gene, the NADH dehydrogenase ND4L
RT and ND5 gene complex, and the glutamyl, methionyl and arginyl tRNA
RT genes from Trichophyton rubrum.";
RL Curr. Genet. 22:229-234(1992).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y18476; CAA77185.1; -
CC F01; S26950; S26950.
CC InterPro; IPR003916; NADHub_oxred5.
CC InterPro; IPR001750; Oxidored_q1.
CC InterPro; IPR001516; Oxidored_q1_N.
CC Pfam; PF00361; oxidored_q1_1_N.
CC Pfam; PF00662; oxidored_q1_N; 1.
CC PRINTS; PR01434; NADHDHGNASE5.
CC OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.
KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 653 AA; 73456 MW; 58207B990F569E14 CRC64;

Query Match 61.4%; Score 54; DB 1; Length 653;
Best Local Similarity 60.0%; Pred. No. 0.7;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPFT 15
DB 185 IIWSFGNLDYSTVFS 199

RESULT 3
ID NUSM PODAN STANDARD; PRT; 652 AA.
AC P20679;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5.
OS Podospora anserina.

Query Match 61.4%; Score 54; DB 1; Length 653;
Best Local Similarity 60.0%; Pred. No. 0.7;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPFT 15
DB 185 IIWSFGNLDYSTVFS 199

RESULT 4
ID NUSM EMBL STANDARD; PRT; 657 AA.
AC P11628; Q00181;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NDHE.
OS Emericella nidulans (Aspergillus nidulans).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VA2 / PYROA4 / CXXC3;
RX MEDLINE=89296483; PubMed=2662141;
RA Brown T.A., Constable A., Ray J.A., Waring R.B., Scazzocchio C.,
RA Davies R.W.;

```

```

OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
OX NCBI_TaxID=5145;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=s, and A;
RX MEDLINE=90204555; PubMed=2319602;
RA Cummings D.J., Michel F., Domenico J.M., McNally K.L.;
RT "DNA sequence analysis of the mitochondrial ND4L-ND5 gene complex
RT from Podospora anserina. Duplication of the ND4L gene within its
RT intron.";
RL J. Mol. Biol. 212:269-286(1990).
RN [2]
RP COMPLETE GENOME.
RC STRAIN=s;
RX MEDLINE=90291512; PubMed=2357736;
RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;
RT "The complete DNA sequence of the mitochondrial genome of Podospora
RT anserina.";
RL Curr. Genet. 17:375-402(1990).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X55026; CAA38798.1; -
CC F01; S09133; S09133.
CC InterPro; IPR003918; NADHub_oxred4.
CC InterPro; IPR003916; NADHub_oxred5.
CC InterPro; IPR001750; Oxidored_q1.
CC InterPro; IPR001516; Oxidored_q1_N.
CC Pfam; PF00361; oxidored_q1_1_N.
CC Pfam; PF00662; oxidored_q1_N; 1.
CC PRINTS; PR01434; NADHDHGNASE5.
CC PRINTS; PR01437; NUOXDRDTASE4.
CC OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.
KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 652 AA; 72632 MW; 84CA887B75755FEB CRC64;

Query Match 58.0%; Score 51; DB 1; Length 652;
Best Local Similarity 53.3%; Pred. No. 2.1;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPFT 15
DB 185 IIWSFGNLDYSTVFS 199

RESULT 4
ID NUSM EMBL STANDARD; PRT; 657 AA.
AC P11628; Q00181;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NDHE.
OS Emericella nidulans (Aspergillus nidulans).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VA2 / PYROA4 / CXXC3;
RX MEDLINE=89296483; PubMed=2662141;
RA Brown T.A., Constable A., Ray J.A., Waring R.B., Scazzocchio C.,
RA Davies R.W.;

```

```

RT "Nucleotide sequence of the Aspergillus nidulans mitochondrial gene
RL for subunit 5 of NADH dehydrogenase.";
RL Nucleic Acids Res. 17:4371-4371(1989).
RN [2]
RNP SEQUENCE OF 316-349 FROM N.A.
RC STRAIN=NRRL 322;
RA Luo X., Khanh N.O., Wientjes F.J.;
RRL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15011; CAA33116.1; -;
CC EMBL; X62993; CAA44727.1; -;
CC PIR; S04724; S04724.
CC InterPro; IPR003916; NADHub oxtreds.
CC InterPro; IPR001750; Oxidored_g1.
CC InterPro; IPR001516; Oxidored_g1_N.
CC Pfam; PF00361; oxidored_g1; 1.
CC Pfam; PF00662; oxidored_g1_N; 1.
CC PRINTS; PR01434; NADHDHGNASE5.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 657 AA; 73361 MW; B5BB679A151E7287 CRC64;
QY Query Match 58.0%; Score 51; DB 1; Length 657;
Db Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VIVSGGNTDYNTPF 14
Db 185 IIVSGGNTDYNTPF 198

```

| | |
|--|---|
| DR | PIR; S64507; S64507. |
| DR | HSSP; P23904; IAJK. |
| DR | GermOnline; I41501; -. |
| DR | SGD; S0003421; CRH1 |
| DR | GO; GO:0009277; C:cell wall (sensu Fungi); IDA. |
| DR | GO; GO:0000131; C:incipient bud site; IDA. |
| DR | InterPro; IPR00875; cona_like lec.gl. |
| DR | InterPro; IPR000757; Glyco_hydro_16. |
| DR | Pfam; PF00722; Glyco_hydro_16; 1. |
| KW | Hypothetical protein. |
| FT | DOMAIN 63 66 POLY-SER. |
| FT | DOMAIN 301 310 POLY-SER. |
| FT | DOMAIN 345 357 POLY-SER. |
| FT | DOMAIN 387 391 POLY-SER. |
| FT | DOMAIN 467 470 POLY-SER. |
| SQ | SEQUENCE 507 AA; 52757 MW; 7D7B61F57AEA942C CRC64; |
| Query Match | |
| Best Local Similarity 57.4%; Score 50.5; DB 1; Length 507; | |
| Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps | |
| QY | 3 WSGGNTDYN-TPFT 15 |
| DB | 233 WAGGETYNDAPFT 246 |
| : : : : : : | |
| RESULT 6 | |
| ARGI HUMAN | |
| ID | ARGI HUMAN STANDARD; PRT; 322 AA. |
| AC | P05089; |
| DC | 13-AUG-1987 (Rel. 05, Created) |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) |
| DE | Arginase 1 (BC 3.5.3.1) (Liver-type arginase). |
| GN | ARG1. |
| OS | Homo sapiens (Human). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| OX | NCBI_TaxID=9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=Liver; |
| RX | MEDLINE=87092419; PubMed=3540966; |
| RA | Haraguchi Y., Takiguchi M., Amaya Y., Kawamoto S., Matsuda I., |
| RA | Mori M.; |
| RT | "Molecular cloning and nucleotide sequence of cDNA for human liver |
| RT | arginase."; |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:412-415(1987). |
| RL | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=Blood; |
| RX | MEDLINE=89016562; PubMed=3174433; |
| RA | Takiguchi M., Haraguchi Y., Mori M.; |
| RT | "Human liver-type arginase gene: structure of the gene and analysis |
| RT | of the promoter region."; |
| RL | Nucleic Acids Res. 16:8789-8802(1988). |

[3] SEQUENCE FROM N.A.
RP Tracey A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RC
RP TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zerborg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gnaratna P.H.,
RA

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smalhus D.E.,
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]
 RN VARIANT ARGININEMIA ARG-235
 RX MEDLINE=9309256; PubMed=1463019;
 RA Uchino T., Haraguchi Y., Aparicio J.M., Mizutani N., Higashikawa M.,
 RA Naitoh H., Mori M., Matsuda I.,
 RT "Three novel mutations in the liver-type arginase gene in three
 RT unrelated Japanese patients with argininemia.";
 RL Am. J. Hum. Genet. 51:1406-1412 (1992).
 [6]
 RN VARIANT SER-290
 RX MEDLINE=92286280; PubMed=1598908;
 RA Grody W.W., Klein D., Dodson A.E., Kern R.M., Wissmann P.B.,
 RA Goodman B.K., Bassand P., Marescau B., Kang S.-S., Leonard J.V.,
 RA Cederbaum S.D.,
 RT "Molecular genetic study of human arginase deficiency.";
 RL Am. J. Hum. Genet. 50:1281-1290 (1992).
 [7]
 RN VARIANTS ARGININEMIA THR-11 AND VAL-138
 RX MEDLINE=95377725; PubMed=7649538;
 RA Uchino T., Snyderman S.E., Lambert M., Qureshi I.A., Shapira S.K.,
 RA Sansaricq C., Smit L.M.E., Jakobs C., Matsuda I.,
 RT "Molecular basis of phenotypic variation in patients with
 RT argininemia.";
 RL Hum. Genet. 96:255-260 (1995).
 CC -!- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
 CC -!- COFACTOR: Manganese.
 CC -!- PATHWAY: Arginine degradation via the urea cycle; first step.
 CC -!- SUBUNIT: Homotrimer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- INDUCTION: By arginine or homocysteine.
 CC -!- DISEASE: Defects in ARG1 are the cause of argininemia
 CC [MIM:207800]. It is characterized by psychomotor retardation and
 CC spastic tetraplegia.
 CC -!- SIMILARITY: Belongs to the arginase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M14502; AAA51776.1; -;
 DR EMBL; X12662; CAA31188.1; -;
 DR EMBL; X12663; CAA31188.1; JOINED.
 DR EMBL; X12664; CAA31188.1; JOINED.
 DR EMBL; X12665; CAA31188.1; JOINED.
 DR EMBL; X12666; CAA31188.1; JOINED.
 DR EMBL; X12667; CAA31188.1; JOINED.
 DR EMBL; X12668; CAA31188.1; JOINED.
 DR EMBL; X12669; CAA31188.1; JOINED.
 DR EMBL; AL121575; CAB92071.1; -;
 DR EMBL; BC020653; AAH20653.1; -;
 DR PIR; S02132; A26370.
 DR HSP; P07824; 3RLA.
 DR Genew; HGNC:663; ARG1.
 DR GK; P05089; -;
 DR MIM; 207800; -;
 DR GO; GO:0005737; Cytoplasm; TAS.
 DR GO; GO:0004053; F:arginase activity; TAS.
 DR GO; GO:0006527; P:arginine catabolism; TAS.

DR InterPro; IPR006035; Arg_agm_form.
 DR InterPro; IPR005924; Arginase.
 DR Pfam; PF00491; arginase; 1.
 DR PRINTS; PRO0116; ARGINASE.
 DR TIGRFAMs; TIGR01229; rocF_arginase; 1.
 DR PROSITE; PS00147; ARGINASE_1; 1.
 DR PROSITE; PS00148; ARGINASE_2; 1.
 DR PROSITE; PS01053; ARGINASE_3; 1.
 KW Urea cycle; Arginine metabolism; Hydrolase; Manganese;
 KW Disease mutation; Polymorphism.
 FT METAL 101 101 MANGANESE 1 (BY SIMILARITY).
 FT METAL 124 124 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 126 126 MANGANESE 2 (BY SIMILARITY).
 FT METAL 128 128 MANGANESE 1 (BY SIMILARITY).
 FT METAL 232 232 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 234 234 MANGANESE 2 (BY SIMILARITY).
 FT VARIANT 11 11 I -> T (in argininemia; 12% of wild-type activity).
 FT FTID=VAR_015594.
 FT VARIANT 138 138 G -> V (in argininemia).
 FT FTID=VAR_015595.
 FT VARIANT 235 235 G -> R (in argininemia).
 FT FTID=VAR_000674.
 FT VARIANT 290 290 T -> S (could be a polymorphism).
 FT FTID=VAR_000675.
 FT CONFLICT 86 86 E -> Q (IN REF. 1).
 SQ SEQUENCE 322 AA; 34735 MW; 8F3BE2652243F622 CRC64;
 Query Match 56.8%; Score 50; DB 1; Length 322;
 Best Local Similarity 60.0%; Pred. No. 1.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VIWGGNTDYNTPT 15
 DB 120 VIWDAHTDINTPLT 134
 RESULT 7
 ARG1 MOUSE
 ID ARG1_MOUSE STANDARD; PRT; 323 AA.
 AC Q61176;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Arginase 1 (EC 3.5.3.1) (Liver-type arginase).
 GN ARG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RA Chieko H.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver, and Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
CC -1- COFACTOR: Manganese.
CC -1- PATHWAY: Arginine degradation via the urea cycle; first step.
CC -1- SUBUNIT: Homotrimer (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the arginase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U51805; AAA98611.1; -
CC EMBL; BC013341; AAH13341.1; -
CC EMBL; BC050005; AAH50005.2; -
CC HSP; P07824; LD3V
CC SWISS-2DPAGE; Q61176; MOUSE.
CC MGD; NGI:88070; Arg1.
CC GO; GO:0004053; F:arginase activity; IDA.
CC InterPro; IPR006035; Arg_agm form.
CC InterPro; IPR005924; Arginase.
CC Pfam; PF00491; arginase; 1.
CC PRINTS; PR00116; ARGINASE.
CC TIGRFAMs; TIGR01229; rocF arginase; 1.
CC PROSITE; PS00147; ARGINASE 1; 1.
CC PROSITE; PS00148; ARGINASE 2; 1.
CC PROSITE; PS01053; ARGINASE 3; 1.
CC Urea cycle; Arginine metabolism; Hydrolase; Manganese.
CC METAL 101 101 MANGANESE 1 (BY SIMILARITY).
CC METAL 124 124 MANGANESE 1 AND 2 (BY SIMILARITY).
CC METAL 126 126 MANGANESE 2 (BY SIMILARITY).
CC METAL 128 128 MANGANESE 1 (BY SIMILARITY).
CC METAL 232 232 MANGANESE 1 AND 2 (BY SIMILARITY).
CC METAL 234 234 MANGANESE 2 (BY SIMILARITY).
CC SEQUENCE 323 AA; 34808 MW; 6D0231978AC1B977 CRC64;
Query Match 56.8%; Score 50; DB 1; Length 323;
Best Local Similarity 60.0%; Pred. No. 1.4;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 VIWGGNTDYNTPPT 15
Db 120 VIWDAHTDINTPLT 134
RESULT 8
ARGI RAT
ID ARG1 RAT STANDARD; PRT; 323 AA.
AC P07824;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Arginase 1 (EC 3.5.3.1) (Liver-type arginase).
GN ARG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=88115364; PubMed=2892837;
RX Ohtake A., Takiguchi M., Shigeto Y., Amaya Y., Kawamoto S., Mori M.,
RA "Structural organization of the gene for rat liver-type arginase.",
RT

RL J. Biol. Chem. 263:2245-2249(1988).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=87194847; PubMed=3571256;
RX Kawamoto S., Amaya Y., Murakami K., Tokunaga F., Iwanaga S.,
RA Kobayashi K., Saeki T., Kimura S., Mori M.,
RT "Complete nucleotide sequence of cDNA and deduced amino acid sequence
of rat liver arginase.",
RL J. Biol. Chem. 262:6280-6283(1987).
[3]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RP MEDLINE=97002331; PubMed=8849731;
RX Kanyo Z.F., Scolnick L.R., Ash D.E., Christianson D.W.;
RT "Structure of a unique binuclear manganese cluster in arginase.",
RL Nature 383:554-557(1996).
[4]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RP MEDLINE=97410344; PubMed=9265637;
RX Scolnick L.R., Kanyo Z.F., Cavalli R.C., Ash D.E., Christianson D.W.;
RT "Altering the binuclear manganese cluster of arginase diminishes
thermostability and catalytic function.",
RL Biochemistry 36:10558-10565(1997).
[5]
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RP PubMed=10542097;
RX Cox J.D., Kim N.N., Traish A.M., Christianson D.W.;
RT "Arginase-boronic acid complex highlights a physiological role in
erectile function.",
RL Nat. Struct. Biol. 6:1043-1047(1999).
CC -1- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
CC -1- COFACTOR: Manganese.
CC -1- PATHWAY: Arginine degradation via the urea cycle; first step.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- INDUCTION: By arginine or homocysteine.
CC -1- SIMILARITY: Belongs to the arginase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17931; AAA40760.1; -
CC EMBL; M17924; AAA40760.1; JOINED.
CC EMBL; M17925; AAA40760.1; JOINED.
CC EMBL; M17926; AAA40760.1; JOINED.
CC EMBL; M17927; AAA40760.1; JOINED.
CC EMBL; M17928; AAA40760.1; JOINED.
CC EMBL; M17929; AAA40760.1; JOINED.
CC EMBL; M17930; AAA40760.1; JOINED.
CC EMBL; J02720; AAA40761.1; -
CC PIR; A26702; A26702.
CC PDB; 1RLA; 15-OCT-97.
CC PDB; 2RLA; 13-MAY-98.
CC PDB; 3RLA; 13-MAY-98.
CC PDB; 4RLA; 13-MAY-98.
CC PDB; 5RLA; 13-MAY-98.
CC PDB; 1D3V; 17-NOV-99.
CC PDB; 1HQ5; 04-APR-01.
CC PDB; 1HQF; 04-APR-01.
CC PDB; 1HGG; 04-APR-01.
CC PDB; 1HCH; 04-APR-01.
CC PDB; 1HQX; 31-DEC-02.
CC InterPro; IPR006035; Arg_agm form.
CC InterPro; IPR005924; Arginase.
CC Pfam; PF00491; arginase; 1.
CC PRINTS; PR00116; ARGINASE.
CC TIGRFAMs; TIGR01229; rocF arginase; 1.
CC PROSITE; PS00147; ARGINASE 1; 1.

DR PROSITE; PS00148; ARGINASE_2; 1.
 DR PROSITE; PS01053; ARGINASE_3; 1.
 KW Urea cycle; Arginine metabolism; Hydrolase; Manganese; 3D-structure.
 FT METAL 101 101 MANGANESE 1.
 FT METAL 124 124 MANGANESE 1 AND 2.
 FT METAL 126 126 MANGANESE 2.
 FT METAL 128 128 MANGANESE 1.
 FT METAL 132 132 MANGANESE 1.
 FT METAL 232 232 MANGANESE 1 AND 2.
 FT METAL 234 234 MANGANESE 2.
 FT CONFLICT 298 298 MANGANESE 1 AND 2.
 FT STRAND 8 12 A -> P (IN REF. 2).
 FT TURN 17 18
 FT HELIX 22 26
 FT HELIX 27 33
 FT TURN 34 35
 FT HELIX 36 40
 FT TURN 41 42
 FT STRAND 47 52
 FT STRAND 64 64
 FT TURN 65 66
 FT STRAND 67 67
 FT HELIX 70 89
 FT TURN 90 91
 FT STRAND 93 97
 FT HELIX 101 103
 FT HELIX 104 114
 FT TURN 116 117
 FT STRAND 119 123
 FT TURN 132 134
 FT HELIX 140 142
 FT STRAND 144 148
 FT STRAND 149 149
 FT HELIX 150 152
 FT TURN 153 154
 FT TURN 160 161
 FT TURN 163 164
 FT STRAND 169 169
 FT HELIX 171 173
 FT STRAND 174 179
 FT HELIX 184 193
 FT TURN 194 194
 FT STRAND 196 199
 FT HELIX 200 206
 FT HELIX 208 220
 FT STRAND 227 232
 FT HELIX 233 235
 FT STRAND 236 236
 FT TURN 238 240
 FT STRAND 252 252
 FT HELIX 254 267
 FT TURN 268 268
 FT STRAND 270 276
 FT TURN 280 281
 FT HELIX 286 303
 FT TURN 304 305
 FT TURN 308 309
 SQ SEQUENCE 323 AA; 34973 MW; 5A92CB0931F9A053 CRC64;

Query Match 56.8%; Score 50; DB 1; Length 323;
 Best Local Similarity 60.0%; Pred. No. 1.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VIWSGGNTDNTPT 15
 DB 120 VIWDAHTDINTPLT 134

RESULT 9
 HV44_MOUSE
 ID HV44_MOUSE STANDARD; PRT; 115 AA.
 AC P01820;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V region P014 precursor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81012133; PubMed=6774258;
 RA Sakano H., Maki R., Kurosawa Y., Roeder W., Tonegawa S.;
 RT "Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes.";
 RL Nature 286:676-683(1980).
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00767; CAA24148.1; -.
 DR PIR; A02095; HVMS14.
 DR PDB; 1A7N; 29-APR-98.
 DR PDB; 1A7O; 29-APR-98.
 DR PDB; 1A7P; 29-APR-98.
 DR PDB; 1A7R; 29-APR-98.
 DR PDB; 1G7H; 17-JAN-01.
 DR PDB; 1G7I; 17-JAN-01.
 DR PDB; 1G7J; 17-JAN-01.
 DR PDB; 1G7N; 17-JAN-01.
 DR PDB; 43C9; 24-JUL-02.
 DR PDB; 43CA; 24-JUL-02.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin V region; Signal; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 115 IG HEAVY CHAIN V REGION P014.
 FT DOMAIN 20 >115 IG-LIKE.
 FT NON TER 115 115
 SQ SEQUENCE 115 AA; 12447 MW; 7569DD4A4843D500 CRC64;

Query Match 53.4%; Score 47; DB 1; Length 115;
 Best Local Similarity 58.3%; Pred. No. 1.4;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDNT 12
 DB 69 MINGDGTSDYNS 80

RESULT 10
 HV43_MOUSE
 ID HV43_MOUSE STANDARD; PRT; 144 AA.
 AC P01819;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig heavy chain V region MOPC 141 precursor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81012133; PubMed=6774258;
 RA Sakano H., Maki R., Kurosawa Y., Roeder W., Tonegawa S.;
 RT "Two types of somatic recombination are necessary for the generation

RT of complete immunoglobulin heavy-chain genes.";
 RL Nature 286:676-683 (1980).
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS TRANSLATED FROM A
 CC DIFFERENTIATED GENE ISOLATED FROM A MYELOMA THAT SECRETES IGG2B.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00768; CAA24149.1; --
 CC PIR; A02094; G2MS14.
 CC HSP; P01825; JFAB.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003596; Ig_v.
 CC Pfam; PF00047; Ig; 1.
 CC SMART; SM00406; IGv; 1.
 CC PROSITE; PS50835; IG LIKE; 1.
 CC Immunoglobulin V region; Signal.
 FT SIGNAL 19
 FT CHAIN 20 144 IG HEAVY CHAIN V REGION MOPC 141.
 FT DOMAIN 20 130 IG-LIKE.
 FT NON TER 144 144
 SQ SEQUENCE 144 AA; 15759 MW; 9B47A7CB3706D30A CRC64;
 Query Match 53.4%; Score 47; DB 1; Length 144;
 Best Local Similarity 63.6%; Pred. No. 1.8; Indels 0; Gaps 0;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 IWSGGNTDYNT 12
 ID IWSM NEUCR STANDARD; PRT; 715 AA.
 AC P05510;
 DT 01-NOV-1988 (Rel. 02, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
 GN ND5.
 OS Neurospora crassa.
 OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=St. Lawrence 74 / SL 74 / ORS 6A;
 RX MEDLINE=8728330; PubMed=3035337;
 RA Nelson M.A., Macino G.;
 RT "Structure and expression of the overlapping ND4L and ND5 genes of
 RT Neurospora crassa mitochondria."
 RL Mol. Gen. Genet. 206:307-317 (1987).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05115; CAB37187.1; --
 CC PIR; S10843; S10843.
 CC InterPro; IPR003918; NADHub_oxred4.

DR InterPro; IPR003916; NADHub_oxred5.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; oxidored_q1; 1.
 DR Pfam; PF00662; oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADHDEGNASES.
 DR PRINTS; PR01437; NUOXDRDTASE4.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 715 AA; 79817 MW; CFAAF0F45FAAC277 CRC64;
 Query Match 52.3%; Score 46; DB 1; Length 715;
 Best Local Similarity 53.3%; Pred. No. 14;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VIWSGGNTDYNTPT 15
 ID VIWSGGNTDYNTPT 15
 DB 185 VLWTLGNLDYATVFS 199
 RESULT 12
 PIG1 DROME
 ID PIG1 DROME STANDARD; PRT; 168 AA.
 AC P26023; O76914; Q24519; Q9W4T3;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 16-OCT-2003 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pre-intermoult gene-1 protein precursor (Gland specific protein).
 GN PIG1 OR PIG-1 OR GSG OR EG:96G10.1 OR CG10790.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Oregon-R, Karsnas, and Samarkand;
 RX MEDLINE=88136816; PubMed=3125018;
 RA Hofmann A., Korge G.;
 RT "Upstream sequences of dosage-compensated and non-compensated alleles
 RT of the larval secretion protein gene Sgs-4 in Drosophila."
 RL Chromosoma 96:1-7 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=92120020; PubMed=1769273;
 RA Furia M., Digilio F.A., Attiaco D., D'Avino P.P., Cavaliere D.,
 RA Polito L.C.;
 RT "Molecular organization of the Drosophila melanogaster Pig-1 gene."
 RL Chromosoma 101:49-54 (1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX Bolshakov V., Borkova D., Minana B., Kafatos F.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Deicher A., Deng Z., Kays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moadry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkian I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: Not known.
 CC -!- TISSUE SPECIFICITY: Low amounts in first to third instar larvae
 CC salivary glands.
 CC -!- DEVELOPMENTAL STAGE: Throughout the larval period.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M24138; AAA28893.1; -;
 DR EMBL; X15760; CAA33767.1; -;
 DR EMBL; AL024484; CAA19669.1; -;
 DR EMBL; AB003427; AAF45859.1; -;
 DR FlyBase; FBgn003086; Pig1.
 KW Signal; Repeat
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 168 PRE-INTERMOULT GENE-1 PROTEIN.
 FT DOMAIN 22 100 ALA/ASP/THR-RICH.
 FT DOMAIN 105 155 ARG/ASN/LYS-RICH.
 FT DOMAIN 27 44 3 X 6 AA TANDEM REPEATS OF S-S-A-D-S-D.
 FT REPEAT 27 32 1.
 FT REPEAT 33 38 2.
 FT REPEAT 33 44 3.
 FT REPEAT 33 44 D -> DSSADSD (IN REF. 2 AND 4).
 FT CONFLICT 38 38 I -> R (IN REF. 1).
 FT CONFLICT 47 47 I -> T (IN REF. 3).
 FT CONFLICT 47 47
 FT CONFLICT 126 168 NNNKKKANNRRKKRRASNNKKKANNRRNNRRNNRR
 FT FT QQQEAWLERKWLKQLHVDVYMHNYKFFCS (IN REF.
 FT FT 4).
 FT FT MISSING (IN REF. 2).
 FT FT KKAANNRRNNRRNNRRRG -> RRHPTTTTGAGTTIPGG
 FT FT VVRAKVVETTTTCGLY (IN REF. 1 AND 2).
 FT FT
 SQ SEQUENCE 168 AA; 18142 MW; 426B6346D6713841 CRC64;
 Query Match 51.1%; Score 45; DB 1; Length 168;
 Best Local Similarity 66.7%; Pred. No. 4.3;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 4 SGSDTDYPTPT 15
 Db |||:|||||
 83 SGSDTDYPTPT 94
 RESULT 13
 ARG2_MOUSE

ID ARG2_MOUSE STANDARD; PRT; 354 AA.
 AC C08691;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Arginase II, mitochondrial precursor (BC 3.5.3.1) (Non-hepatic
 DE arginase) (kidney-type arginase).
 DE ARG2.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RA Iyer R.K.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RL SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RL MEDLINE=99032597; PubMed=9814991;
 EX Morris S.M. Jr., Kepka-Lenhart D., Chen L.C.;
 RA "Differential regulation of arginases and inducible nitric oxide
 RT synthase in murine macrophage cells.";
 RL Am. J. Physiol. 275:E740-E747(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RL MEDLINE=98427201; PubMed=9745037;
 EX Shi O.U., Kepka-Lenhart D., Morris S.M. Jr., O'Brien W.E.;
 RA "Structure of the murine arginase II gene.";
 RL Mamm. Genome 9:822-824(1998).
 CC -!- FUNCTION: May play a role in the regulation of extra-urea cycle
 CC arginine metabolism and also in down-regulation of nitric oxide
 CC synthesis (by similarity).
 CC -!- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
 CC -!- COFACTOR: Manganese.
 CC -!- PATHWAY: Arginine degradation via the urea cycle; first step.
 CC -!- SUBUNIT: Homotrimer (by similarity).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- SIMILARITY: Belongs to the arginase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U90886; AAC22548.1; -;
 DR EMBL; AF032466; AAB86959.1; -;
 DR EMBL; AF045965; AAC78460.1; -;
 DR EMBL; AF044680; AAC78460.1; JOINED.
 DR EMBL; AF045959; AAC78460.1; JOINED.
 DR EMBL; AF045960; AAC78460.1; JOINED.
 DR EMBL; AF045961; AAC78460.1; JOINED.
 DR EMBL; AF045962; AAC78460.1; JOINED.
 DR EMBL; AF045963; AAC78460.1; JOINED.
 DR EMBL; AF045964; AAC78460.1; JOINED.
 DR HSPSP; P07824; ID3V.
 DR MGD; MGI:1330806; Arg2.
 DR InterPro; IPR006035; Arg_agm_form.
 DR InterPro; IPR005924; Arginase.
 DR Pfam; PF00491; arginase; 1.
 DR PRINTS; PR00116; ARGINASE.
 DR TIGRFAMs; TIGR01229; rocF arginase; 1.
 DR PROSITE; PS00147; ARGINASE_1; 1.
 DR PROSITE; PS00148; ARGINASE_2; 1.
 DR PROSITE; PS01053; ARGINASE_3; 1.
 KW Urea cycle; Arginine metabolism; Hydrolase; Manganese;
 KW Transit peptide; Mitochondrion.
 FT TRANSIT 1 22 MITOCHONDRION (POTENTIAL)
 FT

```

FT CHAIN 23 354 ARGINASE II.
FT METAL 120 MANGANESE 1 (BY SIMILARITY).
FT METAL 143 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 145 MANGANESE 2 (BY SIMILARITY).
FT METAL 147 MANGANESE 1 (BY SIMILARITY).
FT METAL 251 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 253 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 354 AA; 38878 MW; B372DF68A19473F2 CRC64;

Query Match 51.1%; Score 45; DB 1; Length 354;
Best Local Similarity 53.3%; Pred. No. 9.3;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNPTPT 15
DB 139 VIWDAHADINTPLT 153

RESULT 14
ARG2_RAT STANDARD; PRT; 354 AA.
AC O08701; P97539;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arginase II, mitochondrial precursor (EC 3.5.3.1) (Non-hepatic
DE arginase) (kidney-type arginase).
GN ARG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RA Iyer R.K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 87-168 FROM N.A.
RC STRAIN=Wistar; TISSUE=Small intestine;
RX MEDLINE=97053663; PubMed=889077;
RA Gotoh T., Sonoki T., Nagasaki A., Terada K., Takiguchi M., Mori M.;
RT "Molecular cloning of cDNA for nonhepatic mitochondrial arginase
RT (arginase II) and comparison of its induction with nitric oxide
RT synthase in a murine macrophage-like cell line.";
RL FEBS Lett. 395:119-122(1996).
CC -!- FUNCTION: May play a role in the regulation of extra-urea cycle
CC arginine metabolism and also in down-regulation of nitric oxide
CC synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
CC -!- COFACTOR: Manganese.
CC -!- PATHWAY: Arginine degradation via the urea cycle; first step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Belongs to the arginase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U90887; AAC22580.1; --
CC DBMEL; D86928; BAA13183.1; --
CC DR HSSP; P07824; 1D3V.
CC DR InterPro; IPR006035; Arg_agm_form.
CC DR InterPro; IPR005924; Arginase.
CC DR Pfam; PF00491; arginase; 1.
CC DR PRINTS; PR00116; ARGINASE.
CC DR TIGRFAMs; TIGR01229; rocF arginase; 1.
CC DR PROSITE; PS00147; ARGINASE_1; 1.

```

```

DR PROSITE; PS00148; ARGINASE 2; 1.
DR PROSITE; PS01053; ARGINASE_3; 1.
KW Urea cycle; Arginine metabolism; Hydrolase; Manganese;
KW Transit peptide; Mitochondrion.
FT TRANSIT 1 22 MITOCHONDRION (POTENTIAL).
FT CHAIN 23 354 ARGINASE II.
FT METAL 120 MANGANESE 1 (BY SIMILARITY).
FT METAL 143 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 145 MANGANESE 2 (BY SIMILARITY).
FT METAL 147 MANGANESE 1 (BY SIMILARITY).
FT METAL 251 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 253 MANGANESE 2 (BY SIMILARITY).
FT CONFLICT 88 Y -> N (IN REF. 2).
FT CONFLICT 90 R -> P (IN REF. 2).
FT CONFLICT 90 A -> S (IN REF. 2).
FT CONFLICT 100 L -> M (IN REF. 2).
FT CONFLICT 116 D -> Y (IN REF. 2).
FT CONFLICT 119 S -> I (IN REF. 2).
FT CONFLICT 128 D -> Y (IN REF. 2).
FT CONFLICT 143 L -> V (IN REF. 2).
FT CONFLICT 164 I -> L (IN REF. 2).
FT CONFLICT 168 I -> L (IN REF. 2).
SQ SEQUENCE 354 AA; 38640 MW; BC03B6BC9B29B8C CRC64;

Query Match 51.1%; Score 45; DB 1; Length 354;
Best Local Similarity 53.3%; Pred. No. 9.3;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNPTPT 15
DB 139 VIWDAHADINTPLT 153

RESULT 15
ARG1_XENLA STANDARD; PRT; 360 AA.
AC Q91553;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 45, Last annotation update)
DE Arginase, non-hepatic 1 (EC 3.5.3.1).
GN ARG1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xeropodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=95014323; PubMed=7929226;
RA Patterson D., Shi Y.-B.;
RT "Thyroid hormone-dependent differential regulation of multiple
RT arginase genes during amphibian metamorphosis.";
RL J. Biol. Chem. 269:25328-25334(1994).
CC -!- FUNCTION: As well as its role in the urea cycle, may be involved
CC in tissue remodeling.
CC -!- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-ornithine + urea.
CC -!- COFACTOR: Manganese (By similarity).
CC -!- PATHWAY: Arginine degradation; first step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed at differing tadpole stages in tail,
CC intestine, hindlimb and trunk region. Most abundant in tadpole
CC tail.
CC -!- DEVELOPMENTAL STAGE: First detected in neurula (stage 16/17).
CC Highest levels in whole tadpole found around stage 47/48. In the
CC intestine, increased levels are found during metamorphosis (stages
CC 58-64). Low levels expressed in hindlimb until stage 66 after
CC which, levels dramatically increase. In the tail, a constant high
CC level of expression is found throughout metamorphosis.
CC -!- INDUCTION: By thyroid hormone (T3).
CC -!- SIMILARITY: Belongs to the arginase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

Thu Oct 14 09:36:58 2004

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U08406; AA56891.1; -
DR PIR; I51663; I51663.
DR HSSP; P07824; 1D3V.
DR InterPro: IPR006035; Arg agm form.
DR InterPro: IPR005924; Arginase.
DR Pfam: PF00491; arginase; 1.
DR PRINTS; PR00116; ARGINASE.
DR TIGRFAMs; TIGR01229; rocG arginase; 1.
DR PROSITE; PS00147; ARGINASE 1; 1.
DR PROSITE; PS00148; ARGINASE 2; 1.
DR PROSITE; PS01053; ARGINASE 3; 1.
KW Urea cycle; Arginine metabolism; Hydrolase; Manganese;
KW Multigene family.
FT METAL 122 122 MANGANESE 1 (BY SIMILARITY).
FT METAL 145 145 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 147 147 MANGANESE 2 (BY SIMILARITY).
FT METAL 149 149 MANGANESE 1 (BY SIMILARITY).
FT METAL 253 253 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 255 255 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 360 AA; 39155 MW; 07B119D8E5F4DA31 CRC64;

Query Match 51.1%; Score 45; DB 1; Length 360;
Best Local Similarity 53.3%; Pred. NO. 9.5;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 VWMSGNTDYNTPFT 15
DB 141 VIWVDAHADINTPLT 155

Search completed: October 6, 2004, 16:30:24
Job time: 11.2105 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004. Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:13 ; Search time 50.2632 seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-4

Perfect score: 88
Sequence: 1 VWSGGNTDYNTEFT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_virus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 53 | 60.2 | 301 | Q9RWL5 | Q9RWL5 deinococcus |
| 2 | 51 | 58.0 | 575 | Q9LTM6 | Q9LTM6 tupaiia herp |
| 3 | 51 | 58.0 | 692 | Q8SHF7 | Q8SHF7 trichoderma |
| 4 | 50 | 56.8 | 236 | Q9BS50 | Q9BS50 homo sapien |
| 5 | 50 | 56.8 | 288 | Q8YXD0 | Q8YXD0 anabaena sp |
| 6 | 50 | 56.8 | 322 | Q95J7C | Q95J7C sus scrofa |
| 7 | 50 | 56.8 | 330 | Q8TE72 | Q8TE72 homo sapien |
| 8 | 50 | 56.8 | 335 | Q8QV14 | Q8QV14 mus musculus |
| 9 | 49 | 55.7 | 482 | Q9LX92 | Q9LX92 mus musculus |
| 10 | 48 | 54.5 | 136 | Q8HHC6 | Q8HHC6 cryptoneutr |
| 11 | 48 | 54.5 | 136 | Q8HAU8 | Q8HAU8 cryptoneutr |
| 12 | 48 | 54.5 | 325 | Q89H72 | Q89H72 bradyrhizob |
| 13 | 48 | 54.5 | 656 | Q8HHD2 | Q8HHD2 cryptoneutr |
| 14 | 48 | 54.5 | 835 | Q97DN6 | Q97DN6 clostidium |
| 15 | 47.5 | 54.0 | 301 | Q872J1 | Q872J1 neurospora |
| 16 | 47 | 53.4 | 692 | Q97YM6 | Q97YM6 sulfolobus |

ALIGNMENTS

RESULT 1

Q9RWL5
ID Q9RWL5 PRELIMINARY; PRT; 301 AA.
AC Q9RWL5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Arginase.
GN DR0651.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20038896; PubMed=10567266; F.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RT Science 286:1571-1577(1999).
RL EMBL; AE001922; AAF10226.1; .
DR PIR: H75493; H75493.
DR HSSP; P51608; 2CEV.
DR TIGR; DR0651; .
DR GO; GO:0004053; F:arginase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0006527; P:arginine catabolism; IEA.
DR InterPro; IPR005924; Arginase.
DR InterPro; IPR006035; Arg agm_form.
DR Pfam; PF00491; arginase; .
DR PRINTS; PR00116; ARGINASE.
DR TIGRFAMs; TIGR01229; rocF_arginase; 1.

Q826K9 streptomyc
Q91J28 streptomyc
Q8KJF5 candida ole
Q97M66 clostridium
Q9R1G7 mus musculu
P73589 synecocyst
Q897X5 clostridium
Q8ZTQ2 pyrobaculum
Q8U3G3 pyrococcus
Q58339 pyrococcus
P74734 synecocyst
Q9UYU1 pyrococcus
Q9AA20 caulobacter
Q56785 staphylococ
Q8HG56 verticilliu
Q8AAJ2 bacteroides
Q42800 aspergillus
Q9J0P4 aspergillus
Q7Y4Q7 bacterioph
Q92W51 rhizobium m
Q8E147 shewanella
Q42795 collettotric
Q39012 trichoderma
Q9NB58 leishmania
Q8DAN2 vibrio vuln
Q8KJF5 streptococc
Q89HH8 bradyrhizob
Q99NG4 mus musculu
Q8ML14 drosophila

17 46 52.3 199 16 Q826K9
18 46 52.3 201 16 Q9R1J28
19 46 52.3 425 3 Q8NK99
20 45 51.1 319 16 Q97M66
21 45 51.1 354 11 Q9R1G7
22 45 51.1 369 16 P73589
23 45 51.1 423 16 Q897X5
24 45 51.1 434 17 Q8ZTQ2
25 45 51.1 448 17 Q8U3G3
26 45 51.1 457 17 Q58339
27 45 51.1 470 16 P74734
28 45 51.1 480 17 Q9UYU1
29 45 51.1 549 16 Q9AA20
30 45 51.1 628 9 Q56785
31 45 51.1 642 8 Q8HG56
32 44.5 50.6 252 16 Q8AAJ2
33 44.5 50.6 302 3 Q42800
34 44.5 50.6 395 3 Q8U0P4
35 44 50.0 257 9 Q7Y4Q7
36 44 50.0 308 16 Q92W51
37 44 50.0 577 16 Q8E147
38 44 50.0 600 3 Q42795
39 44 50.0 662 3 Q39012
40 44 50.0 682 5 Q9NB58
41 44 50.0 926 16 Q8DAN2
42 44 50.0 1146 16 Q8KJF5
43 44 50.0 1195 16 Q89HH8
44 43 48.9 121 11 Q99NG4
45 43 48.9 130 5 Q8ML14

```

DR PROSITE; PS00147; ARGINASE_1; 1.
DR PROSITE; PS00148; ARGINASE_2; 1.
DR KW Complete proteome.
SQ SEQUENCE 301 AA; 32339 MW; 9DD18F99E744682 CRC64;

Query Match 60.2%; Score 53; DB 16; Length 301;
Best Local Similarity 69.2%; Pred. No. 1.9;
Matches 9; Conservative 1; Mismatches 0; Indels 3; Gaps 0;

QY 1 VWSGGNTDYNTP 13
Db 118 VWSGGNTDYNTP 130

RESULT 2
Q91TN6 PRELIMINARY; PRT; 575 AA.
AC Q91TN6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE T52.
OS Tupaia herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Retaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RA Bahr U., Darai G.;
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree
Shrew) Herpesvirus."
RL J. Virol. 75:4854-4870 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Darai G., Bahr U.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281817; AAK5101.1; -.
DR GO; GO:0019031; Civil viral envelope; IEA.
DR InterPro; IPR002597; Herpes_env.
DR Pfam; PF01673; Herpes_env; 1.
SQ SEQUENCE 575 AA; 63743 MW; BE62EEF35166825D CRC64;

Query Match 58.0%; Score 51; DB 12; Length 575;
Best Local Similarity 57.1%; Pred. No. 8.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IWSGGNTDYNTP 15
Db 302 VWSGGNTDYNTP 315

RESULT 3
Q8SHP7 PRELIMINARY; PRT; 692 AA.
AC Q8SHP7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5.
GN ND5.
OS Trichoderma reesei (Hypocrea jecorina).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21950703; PubMed=11825887;
RA Chambergo F.S., Bonaccorsi E.D., Ferreira A.J.S., Ramos A.S.P.,
Ferreira Junior J.R., Abrahao-Neto J., Farah J.P.S., El-Dorry H.;

```

```

RT "Elucidation of the Metabolic Fate of Glucose in the Filamentous
Fungus Trichoderma reesei Using Expressed Sequence Tag (EST) Analysis
and cDNA Microarrays."
RL J. Biol. Chem. 277:13983-13988 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Chambergo F.S., Bonaccorsi E.D., Ferreira A.J.S., Ramos A.S.P.,
Ferreira Junior J.R., Abrahao-Neto J., Farah J.P.S., El-Dorry H.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF447590; AAL74164.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR003916; NADH_oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF00662; oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHDHGNASES.
DR NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
SQ SEQUENCE 692 AA; 77171 MW; 89821FAAFCSA6C81 CRC64;

Query Match 58.0%; Score 51; DB 8; Length 692;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTP 15
Db 185 ILWSGLNLDYSTVFS 199

RESULT 4
Q9BS50 PRELIMINARY; PRT; 236 AA.
AC Q9BS50;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to arginase, liver.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005321; AAH05321.1; -.
DR EMBL; BT006741; AAP35387.1; -.
DR HSP; P07824; 3RLA.
DR GO; GO:0004053; F:arginase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0006527; P:arginine catabolism; IEA.
DR InterPro; IPR005924; Arginase.
DR InterPro; IPR006035; Arg_agn_form.
DR Pfam; PF00491; arginase; 1.
DR PRINTS; PR00116; ARGINASE.
DR TIGRfams; TIGR01229; rcf arginase; 1.
DR PROSITE; PS00147; ARGINASE_1; 1.
DR PROSITE; PS00148; ARGINASE_2; 1.
SQ SEQUENCE 236 AA; 25356 MW; 79F02C69B700AB67 CRC64;

Query Match 56.8%; Score 50; DB 4; Length 236;

```

Best Local Similarity 60.0%; Pred. No. 4.5;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 VWSGGNTDYNTPFT 15
Db 120 VIWVDAHTDINTPLT 134

RESULT 5
Q8YXD0 PRELIMINARY; PRT; 288 AA.
AC Q8YXD0;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Permease protein of ABC transporter.
GN ALL1284.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2159285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001)
DR EMBL: AF003585; BAB73241.1; -;
DR PIR: A11966; A11966.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; F:transport activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR001851; Bac_inmem_transp.
DR Pfam: PF02653; BPD_transp_2; 1.
KW Complete proteome.
SQ SEQUENCE 288 AA; 30913 MW; AE7C2E0D454E41FC CRC64;

Query Match 56.8%; Score 50; DB 16; Length 288;
Best Local Similarity 53.3%; Pred. No. 5.7;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 VWSGGNTDYNTPFT 15
Db 114 LIWGRNQYNLPIT 128

RESULT 6
Q95JC8 PRELIMINARY; PRT; 322 AA.
ID Q95JC8;
AC Q95JC8;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Arginase I (EC 3.5.3.1).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kepka-Lenhart D., Morris S.M. Jr.;
RT "Sequence of Pig Arginase I";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK091874.1; -;
DR GO: GO:0004053; F:arginase activity; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0006527; P:arginine catabolism; IEA.
DR InterPro: IPR005924; Arginase.
DR InterPro: IPR006035; Arg_agm_form.

Pfam: PF00491; arginase; 1.
DR PRINTS: PR00116; ARGINASE.
DR TIGRFAMs: TIGR01229; rocF arginase; 1.
DR PROSITE: PS00147; ARGINASE_1; 1.
DR PROSITE: PS00148; ARGINASE_2; 1.
DR PROSITE: PS01053; ARGINASE_3; 1.
KW Hydrolase.
SQ SEQUENCE 322 AA; 35018 MW; D99FB071CE916B2F CRC64;

Query Match 56.8%; Score 50; DB 6; Length 322;
Best Local Similarity 60.0%; Pred. No. 6.4;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 VWSGGNTDYNTPFT 15
Db 120 VIWVDAHTDINTPLT 134

RESULT 7
Q8TE72 PRELIMINARY; PRT; 330 AA.
ID Q8TE72;
AC Q8TE72;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Arginase type I erythroid variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee Y.T., Miller J.L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY074488; AAL71547.1; -;
DR GO: GO:0004053; F:arginase activity; IEA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0006527; P:arginine catabolism; IEA.
DR InterPro: IPR005924; Arginase.
DR InterPro: IPR006035; Arg_agm_form.
DR Pfam: PF00491; arginase; 1.
DR PRINTS: PR00116; ARGINASE.
DR TIGRFAMs: TIGR01229; rocF arginase; 1.
DR PROSITE: PS00147; ARGINASE_1; 1.
DR PROSITE: PS00148; ARGINASE_2; 1.
DR PROSITE: PS01053; ARGINASE_3; 1.
SQ SEQUENCE 330 AA; 35664 MW; 5747D414E8741D7B CRC64;

Query Match 56.8%; Score 50; DB 4; Length 330;
Best Local Similarity 60.0%; Pred. No. 6.6;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 VWSGGNTDYNTPFT 15
Db 128 VIWVDAHTDINTPLT 142

RESULT 8
Q80V14 PRELIMINARY; PRT; 335 AA.
ID Q80V14;
AC Q80V14;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Arginase 1, liver.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZ5CH II; TISSUE=Breast tumor;
RA Strausberg R.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC050005; AAH50005.1; -
 DR GO; GO:0004053; F:arginase activity; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0006527; P:arginine catabolism; IEA.
 DR InterPro; IPR005924; Arginase.
 DR InterPro; IPR006035; Arg_agn_form.
 DR Pfam; PF00491; arginase; 1.
 DR PRINTS; PR00116; ARGINASE.
 DR TIGRFAMs; TIGR01229; rocF arginase; 1.
 DR PROSITE; PS00147; ARGINASE_1; 1.
 DR PROSITE; PS00148; ARGINASE_2; 1.
 DR PROSITE; PS01053; ARGINASE_3; 1.
 DR PROSITE; PS01053; ARGINASE_3; 1.
 SQ SEQUENCE 335 AA; 36260 MW; F32D0A2CA835A07D CRC64;

Query Match 56.8%; Score 50; DB 11; Length 335;
 Best Local Similarity 60.0%; Pred. No. 6.7;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPT 15
 :|||:|||
 Db 132 VIWVDAHTDINTPLT 146

RESULT 9

Q91X92 PRELIMINARY; PRT; 482 AA.
 AC Q91X92;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC011181; AAH11181.1; -
 DR PIR; F33932; F33932.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 4.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS0835; IG LIKE; 4.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Hypothetical protein
 SQ SEQUENCE 482 AA; 51865 MW; 312E01F9C1BC7F3C CRC64;

Query Match 55.7%; Score 49; DB 11; Length 482;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNT 12
 :|||:|||
 Db 69 VIWGTGVNTYNS 80

RESULT 10

Q8HHC6 PRELIMINARY; PRT; 136 AA.

AC Q8HHC6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NADH-ubiquinone oxidoreductase subunit 5 (fragment).
 GN ND5.
 OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia parasitica).

OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Diaporthales; Valsaceae;
 OC Cryphonectria-Endothia complex; Cryphonectria.
 OX NCBI_TaxID=5116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=I-7;
 RA Gobbi E.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF456841; AA014107.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR003916; NADH_oxred5.
 DR InterPro; IPR001750; Oxidored_g1.
 DR Pfam; PF00361; Oxidored_g1; 1.
 DR Pfam; PF00662; Oxidored_g1_N; 1.
 DR PRINTS; PR01434; NADHGNASES.
 DR Ubiquinone; Mitochondrion.
 KW Ubiquinone; Mitochondrion.
 FT NON_TER 1
 FT NON_TER 136
 SQ SEQUENCE 136 AA; 15244 MW; F9ACEEBE138F3D99 CRC64;

Query Match 54.5%; Score 48; DB 8; Length 136;
 Best Local Similarity 53.3%; Pred. No. 5.2;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPT 15
 :|||:|||
 Db 102 ILWSLGNLDYATVFS 116

RESULT 11

Q8HAUS PRELIMINARY; PRT; 136 AA.
 AC Q8HAUS;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NADH-ubiquinone oxidoreductase subunit 5 (fragment).
 GN ND5.
 OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia parasitica).
 OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Diaporthales; Valsaceae;
 OC Cryphonectria-Endothia complex; Cryphonectria.
 OX NCBI_TaxID=5116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=I-3, I-4, I-9, I-10, and I-12b;
 RA Gobbi E.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF456839; AA014104.1; -
 DR EMBL; AF456840; AA014106.1; -
 DR EMBL; AF456842; AA014108.1; -
 DR EMBL; AF456843; AA014109.1; -
 DR EMBL; AF456844; AA014111.1; -
 DR EMBL; AF456845; AA014114.1; -
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR003916; NADH_oxred5.
 DR InterPro; IPR001750; Oxidored_g1.
 DR Pfam; PF00361; Oxidored_g1; 1.
 DR Pfam; PF00662; Oxidored_g1_N; 1.
 DR PRINTS; PR01434; NADHGNASES.
 KW Ubiquinone; Mitochondrion.
 FT NON_TER 1
 FT NON_TER 136

SQ SEQUENCE 136 AA; 15204 MW; F9B296B65EB99BD99 CRC64;
Query Match 54.5%; Score 48; DB 8; Length 136;
Best Local Similarity 53.3%; Pred. No. 5.2;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 VWSGGNTDYNTPPT 15
Db 102 ILWSLGNLDYATVFS 116
RESULT 12
Q89H72 ID Q89H72 PRELIMINARY; PRT; 325 AA.
AC Q89H72
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Arginase.
GN BL6122.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BAC51387.1; -;
DR GO; GO:0004053; F:arginase activity; IEA.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0006527; P:arginine catabolism; IEA.
DR InterPro; IPR005924; Arginase.
DR InterPro; IPR006035; Arg_agm_form.
DR Pfam; PF00491; arginase; 1.
DR PRINTS; PR00116; ARGINASE.
DR PROSITE; PS00147; ARGINASE_1; 1.
DR PROSITE; PS00148; ARGINASE_2; 1.
DR PROSITE; PS01053; ARGINASE_3; 1.
KW Complete proteome.
SQ SEQUENCE 325 AA; 35660 MW; 92964D631B8D2BCD CRC64;
Query Match 54.5%; Score 48; DB 16; Length 325;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 VWSGGNTDYNTPPT 15
Db 131 VLMWDADYNTPT 145
RESULT 13
Q89HD2 ID Q89HD2 PRELIMINARY; PRT; 656 AA.
AC Q89HD2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH-ubiquinone oxidoreductase subunit 5.
GN ND5.
OS Cryphonectria parasitica (Chesnut blight fungus) (Endothia
OS Parasitica).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Diaporthales; Valsaceae;
OC Cryphonectria-Endothia complex; Cryphonectria.

OX NCBI_TaxID=5116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ep 155;
RA Gobbi E.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF456838; AA014099.1; -;
DR GO; GO:0005739; F:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR003918; NADHub_oxred4.
DR InterPro; IPR003916; NADHub_oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF00662; oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHGRNASE5.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW Ubiquinone; Mitochondrion.
SQ SEQUENCE 656 AA; 72600 MW; 35C99E836E257DEA CRC64;
Query Match 54.5%; Score 48; DB 8; Length 656;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 VWSGGNTDYNTPPT 15
Db 185 ILWSLGNLDYATVFS 199
RESULT 14
Q97DN6 ID Q97DN6 PRELIMINARY; PRT; 835 AA.
AC Q97DN6
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable alpha-arabinofuranosidase.
GN CAC3436.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286; M.V., Makarova K.S., Zeng Q.,
RA Noelling J., Breton G., Omelchenko M.V., Gou D., Hitti J., Wolf Y.I.,
RA Gibson R., Lee H.M., Dubois J., Gou D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007841; AAK81386.1; -;
DR FIR; C97322; C97322.
DR InterPro; IPR003305; CBM_CenC.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF02018; CBM_4_9; 1.
KW Complete proteome.
SQ SEQUENCE 835 AA; 93561 MW; 35C11C69D587CEFC CRC64;
Query Match 54.5%; Score 48; DB 16; Length 835;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 VWSGGNTDYNTPPT 13
Db 520 MIFNGNTNVTPT 532
RESULT 15
Q872J1

```

ID Q872J1 PRELIMINARY; PRT; 301 AA.
AC Q872J1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Probable rAsp f 9 allergen.
GN 64C2.250.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX294009; CAD70740.1; -.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR008985; ConA-like lec.gl.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 301 AA; 31771 MW; FCB56E1498A800C2 CRC64;

Query Match 54.0%; Score 47.5; DB 3; Length 301;
Best Local Similarity 56.2%; Pred. No. 15;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 VIWGGNTDYNT-PPT 15
Db : ||| |||
229 IAWAGGVTDYTRGPFT 244

```

Search completed: October 6, 2004, 16:33:47
 Job time : 55.2832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 93.4211 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-4

Perfect score: 88

Sequence: 1 VWSGGNTDYNTPT 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 segs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------------|
| 1 | 88 | 100.0 | 16 | 10 | US-09-798-689-28 |
| 2 | 88 | 100.0 | 16 | 10 | US-09-996-954B-4 |
| 3 | 88 | 100.0 | 16 | 12 | US-10-374-600-3 |
| 4 | 88 | 100.0 | 16 | 15 | US-10-374-531-3 |
| 5 | 88 | 100.0 | 18 | 12 | US-10-374-600-2 |
| 6 | 88 | 100.0 | 18 | 15 | US-10-374-531-2 |
| 7 | 88 | 100.0 | 119 | 12 | US-10-374-600-115 |
| 8 | 88 | 100.0 | 119 | 12 | US-10-374-600-116 |
| 9 | 88 | 100.0 | 119 | 12 | US-10-374-600-117 |
| 10 | 88 | 100.0 | 119 | 12 | US-10-374-600-118 |
| 11 | 88 | 100.0 | 119 | 12 | US-10-374-600-119 |
| 12 | 88 | 100.0 | 119 | 12 | US-10-374-600-120 |
| 13 | 88 | 100.0 | 119 | 15 | US-10-374-531-115 |
| 14 | 88 | 100.0 | 119 | 15 | US-10-374-531-116 |
| 15 | 88 | 100.0 | 119 | 15 | US-10-374-531-117 |

| | | | | | | |
|----|------|-------|-----|----|-------------------|-------------------|
| 16 | 88 | 100.0 | 119 | 15 | US-10-374-531-118 | Sequence 118, App |
| 17 | 88 | 100.0 | 119 | 15 | US-10-374-531-119 | Sequence 119, App |
| 18 | 88 | 100.0 | 119 | 15 | US-10-374-531-120 | Sequence 120, App |
| 19 | 88 | 100.0 | 138 | 12 | US-10-374-600-8 | Sequence 8, Appli |
| 20 | 88 | 100.0 | 138 | 12 | US-10-374-600-13 | Sequence 13, Appl |
| 21 | 88 | 100.0 | 138 | 12 | US-10-374-600-19 | Sequence 19, Appl |
| 22 | 88 | 100.0 | 138 | 15 | US-10-374-531-8 | Sequence 8, Appli |
| 23 | 88 | 100.0 | 138 | 15 | US-10-374-531-13 | Sequence 13, Appl |
| 24 | 88 | 100.0 | 138 | 15 | US-10-374-531-19 | Sequence 19, Appl |
| 25 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-24 | Sequence 24, Appl |
| 26 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-26 | Sequence 26, Appl |
| 27 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-27 | Sequence 27, Appl |
| 28 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-28 | Sequence 28, Appl |
| 29 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-29 | Sequence 29, Appl |
| 30 | 76.5 | 86.9 | 132 | 12 | US-10-374-600-30 | Sequence 30, Appl |
| 31 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-24 | Sequence 24, Appl |
| 32 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-26 | Sequence 26, Appl |
| 33 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-27 | Sequence 27, Appl |
| 34 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-28 | Sequence 28, Appl |
| 35 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-29 | Sequence 29, Appl |
| 36 | 76.5 | 86.9 | 132 | 15 | US-10-374-531-30 | Sequence 30, Appl |
| 37 | 70 | 79.5 | 116 | 9 | US-09-865-483-13 | Sequence 13, Appl |
| 38 | 70 | 79.5 | 114 | 9 | US-09-865-483-8 | Sequence 8, Appl |
| 39 | 65 | 73.9 | 16 | 12 | US-10-239-656-35 | Sequence 35, Appl |
| 40 | 65 | 73.9 | 16 | 12 | US-10-239-656-45 | Sequence 45, Appl |
| 41 | 65 | 73.9 | 116 | 16 | US-10-682-845-53 | Sequence 53, Appl |
| 42 | 65 | 73.9 | 119 | 12 | US-10-389-417-19 | Sequence 19, Appl |
| 43 | 65 | 73.9 | 119 | 12 | US-10-389-417-20 | Sequence 20, Appl |
| 44 | 65 | 73.9 | 119 | 12 | US-10-452-357-37 | Sequence 37, Appl |
| 45 | 65 | 73.9 | 119 | 12 | US-10-452-357-60 | Sequence 60, Appl |

ALIGNMENTS

RESULT 1

US-09-798-689-28
; Sequence 28, Application US/09798689
; Publication No. US20030103973A1
; GENERAL INFORMATION:
; APPLICANT: Rockwell, Patricia
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: Method for Reducing Tumor Growth with VEGF Antagonists
; TITLE OF INVENTION: Combined With Radiation and Chemotherapy
; FILE REFERENCE: Sequence Listings 1-41 for 381-25 CIP
; CURRENT APPLICATION NUMBER: US/09/798.689
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/401,163
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/967,113
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 08/706,804
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/476,533
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/326,552
; PRIOR FILING DATE: 1994-10-20
; PRIOR APPLICATION NUMBER: 08/196,041
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Mouse
US-09-798-689-28

Query Match 100.0%; Score 88; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 VWSGGNTDYNTPT 15
|||||

Db 1 VIWSGGNTDYNTPFT 15

RESULT 2

US-09-996-954B-4

Sequence 4, Application US/09996954B

Publication No. US20030157104A1

GENERAL INFORMATION:

APPLICANT: Waksal, Harlan W.

TITLE OF INVENTION: Treatment of Refractory Human Tumors

TITLE OF INVENTION: With Epidermal Growth Factor Receptor Antagonists

FILE REFERENCE: 11245-46605

CURRENT APPLICATION NUMBER: US/09/996,954B

CURRENT FILING DATE: 2001-11-30

PRIOR APPLICATION NUMBER: 09/840,146

PRIOR FILING DATE: 04-24-2001

PRIOR APPLICATION NUMBER: 09/374,028

PRIOR FILING DATE: 08-13-1999

PRIOR APPLICATION NUMBER: 09/312,284

PRIOR FILING DATE: 05-14-1999

NUMBER OF SEQ ID NOS: 12

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 16

TYPE: PRT

ORGANISM: Homo sapiens-Rodent Chimera

US-09-996-954B-4

Query Match 100.0%; Score 88; DB 10; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15

Db 1 VIWSGGNTDYNTPFT 15

RESULT 3

US-10-374-600-3

Sequence 3, Application US/10374600

Publication No. US20030224001A1

GENERAL INFORMATION:

APPLICANT: InClone Systems Incorporated, et al.

TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,600

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

US-09-996-954B-4

Query Match 100.0%; Score 88; DB 12; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15

Db 1 VIWSGGNTDYNTPFT 15

RESULT 4

US-10-374-531-3

Sequence 3, Application US/10374531

Publication No. US20040006212A1

GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.

TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,531

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

US-10-374-600-3

Query Match 100.0%; Score 88; DB 12; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.4e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15

Db 1 VIWSGGNTDYNTPFT 15

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

```
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-374-531-3

Query Match      100.0%; Score 88; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15
Db 1 VIWSGGNTDYNTPFT 15

RESULT 5
US-10-374-600-2
; Sequence 2, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-374-600-2

Query Match      100.0%; Score 88; DB 12; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15
Db 1 VIWSGGNTDYNTPFT 15

US-10-374-600-2

Query Match      100.0%; Score 88; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15
Db 1 VIWSGGNTDYNTPFT 15

RESULT 6
US-10-374-531-2
; Sequence 2, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-374-531-2

Query Match      100.0%; Score 88; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPFT 15
Db 2 VIWSGGNTDYNTPFT 16

RESULT 7
US-10-374-600-115
; Sequence 115, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
```

INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSER: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 115:

US-10-374-600-115

Query Match 100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
DB 50 VWSGGNTDYNTPFT 64

RESULT 8

US-10-374-600-116
Sequence 116, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 116:

US-10-374-600-116

Query Match 100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
DB 50 VWSGGNTDYNTPFT 64

RESULT 9

US-10-374-600-117
Sequence 117, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 117:

US-10-374-600-117

Query Match 100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
DB 50 VWSGGNTDYNTPFT 64

COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 116:

US-10-374-600-116

Query Match 100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
DB 50 VWSGGNTDYNTPFT 64

RESULT 9

US-10-374-600-117
Sequence 117, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
SEQUENCE DESCRIPTION: SEQ ID NO: 117:

US-10-374-600-117

Query Match 100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
DB 50 VWSGGNTDYNTPFT 64

```

; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
;   NAME: Deborah A. Somerville
;   REGISTRATION NUMBER: 31,995
;   REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 425-7200
; INFORMATION FOR SEQ ID NO: 117:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 119 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;     TOPOLOGY: not relevant
;     MOLECULE TYPE: peptide
;     HYPOTHETICAL: NO
;     FRAGMENT TYPE: internal
;   SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-10-374-600-117

Query Match          100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
Db 50 VWSGGNTDYNTPFT 64

RESULT 10
US-10-374-600-118
; Sequence 118, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; INFORMATION FOR SEQ ID NO: 119:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 119 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;     TOPOLOGY: not relevant
;     MOLECULE TYPE: peptide
;     HYPOTHETICAL: NO
;     FRAGMENT TYPE: internal
;   SEQUENCE DESCRIPTION: SEQ ID NO: 119:
US-10-374-600-119

Query Match          100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPFT 15
Db 50 VWSGGNTDYNTPFT 64

RESULT 11
US-10-374-600-119
; Sequence 119, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; INFORMATION FOR SEQ ID NO: 119:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 119 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;     TOPOLOGY: not relevant
;     MOLECULE TYPE: peptide
;     HYPOTHETICAL: NO
;     FRAGMENT TYPE: internal
;   SEQUENCE DESCRIPTION: SEQ ID NO: 119:
US-10-374-600-119
```

```
US-10-374-600-119
Query Match      100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPFT 15
   |||||
Db 50 VIWGGNTDYNTPFT 64

RESULT 12
US-10-374-600-120
; Sequence 120, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 120:
1 VIWGGNTDYNTPFT 15
|||||
50 VIWGGNTDYNTPFT 64

US-10-374-600-120
Query Match      100.0%; Score 88; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPFT 15
   |||||
Db 50 VIWGGNTDYNTPFT 64

US-10-374-531-115
; Sequence 115, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 115:
1 VIWGGNTDYNTPFT 15
|||||
50 VIWGGNTDYNTPFT 64

US-10-374-531-116
; Sequence 116, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
```

```
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-7200
TELEFAX: (212) 425-5288
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-10-374-531-116
Query Match 100.0%; Score 88; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VIWSGGNTDYNTPFT 15
Db 50 VIWSGGNTDYNTPFT 64
RESULT 15
US-10-374-531-117
; Sequence 117, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-7200
TELEFAX: (212) 425-5288
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-10-374-531-117
Query Match 100.0%; Score 88; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VIWSGGNTDYNTPFT 15
Db 50 VIWSGGNTDYNTPFT 64
Search completed: October 6, 2004, 17:09:03
Job time : 94.4211 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 26.5789 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-4
Perfect score: 88
Sequence: 1 VWSGGNTDYNTPF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 70 | 79.5 | 16 | 4 | US-09-865-483-13 |
| 2 | 70 | 79.5 | 114 | 4 | US-09-865-483-8 |
| 3 | 65 | 73.9 | 119 | 1 | US-07-634-278-37 |
| 4 | 65 | 73.9 | 119 | 1 | US-07-634-278-60 |
| 5 | 65 | 73.9 | 119 | 1 | US-07-634-278-61 |
| 6 | 65 | 73.9 | 119 | 1 | US-08-477-728-37 |
| 7 | 65 | 73.9 | 119 | 1 | US-08-477-728-60 |
| 8 | 65 | 73.9 | 119 | 1 | US-08-477-728-61 |
| 9 | 65 | 73.9 | 119 | 1 | US-08-474-040-37 |
| 10 | 65 | 73.9 | 119 | 1 | US-08-474-040-60 |
| 11 | 65 | 73.9 | 119 | 1 | US-08-474-040-61 |
| 12 | 65 | 73.9 | 119 | 1 | US-08-487-200-37 |
| 13 | 65 | 73.9 | 119 | 1 | US-08-487-200-60 |
| 14 | 65 | 73.9 | 119 | 1 | US-08-487-200-61 |
| 15 | 65 | 73.9 | 119 | 3 | US-08-484-537-37 |
| 16 | 65 | 73.9 | 119 | 3 | US-08-484-537-60 |
| 17 | 65 | 73.9 | 119 | 3 | US-08-484-537-61 |
| 18 | 65 | 73.9 | 138 | 1 | US-07-634-278-33 |
| 19 | 65 | 73.9 | 138 | 1 | US-08-477-728-33 |
| 20 | 65 | 73.9 | 138 | 1 | US-08-474-040-33 |
| 21 | 65 | 73.9 | 138 | 1 | US-08-487-200-33 |
| 22 | 65 | 73.9 | 138 | 3 | US-08-484-537-33 |
| 23 | 60 | 68.2 | 16 | 1 | US-08-432-694-20 |
| 24 | 59 | 67.0 | 264 | 3 | US-08-564-164A-4 |
| 25 | 58 | 65.9 | 242 | 6 | 5455030-15 |
| 26 | 54 | 61.4 | 116 | 3 | US-08-397-411-3 |
| 27 | 54 | 61.4 | 116 | 3 | US-08-397-411-4 |

28 54 61.4 273 3 US-08-397-411-6 Sequence 6, Appli
29 54 61.4 446 3 US-08-397-411-7 Sequence 7, Appli
30 53 60.2 97 3 US-08-881-037-66 Sequence 66, Appli
31 53 60.2 112 4 US-09-189-129-3 Sequence 3, Appli
32 53 60.2 116 2 US-08-308-494A-21 Sequence 21, Appli
33 53 60.2 121 3 US-08-881-037-67 Sequence 67, Appli
34 53 60.2 222 2 US-08-190-199A-67 Sequence 67, Appli
35 53 60.2 235 2 US-08-190-199A-61 Sequence 61, Appli
36 52 59.1 31 4 US-08-525-539A-36 Sequence 36, Appli
37 52 59.1 107 1 US-07-942-245-14 Sequence 14, Appli
38 52 59.1 111 2 US-08-470-139-6 Sequence 6, Appli
39 52 59.1 111 4 US-09-347-061-6 Sequence 6, Appli
40 52 59.1 111 4 US-09-537-911A-67 Sequence 67, Appli
41 52 59.1 120 2 US-08-353-372A-36 Sequence 36, Appli
42 52 59.1 120 2 US-08-652-558-4 Sequence 4, Appli
43 52 59.1 120 2 US-08-652-558-5 Sequence 5, Appli
44 52 59.1 120 2 US-08-652-558-6 Sequence 6, Appli
45 52 59.1 120 2 US-08-652-558-7 Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-865-483-13
; Sequence 13, Application US/09865483
; Patent No. 6680053
; GENERAL INFORMATION:
; APPLICANT: LEE, Jong Wook et al.
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV S-SU
; FILE REFERENCE: 1599-0197P
; CURRENT APPLICATION NUMBER: US/09/865,483
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-865-483-13

Query Match 79.5%; Score 70; DB 4; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.00045;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VWSGGNTDYNTPF 14
Db 1 VWSGGNTDYNAAF 14
|||||||

RESULT 2
US-09-865-483-8
; Sequence 8, Application US/09865483
; Patent No. 6680053
; GENERAL INFORMATION:
; APPLICANT: LEE, Jong Wook et al.
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV S-SU
; FILE REFERENCE: 1599-0197P
; CURRENT APPLICATION NUMBER: US/09/865,483
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-865-483-8

Query Match 79.5%; Score 70; DB 4; Length 114;
Best Local Similarity 85.7%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWGGTNDYNTPF 14
Db 50 VIWGGTNDYNAAP 63

RESULT 4
US-07-634-278-60
; Sequence 60, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION NUMBER: US 07/310,252
; APPLICATION DATA:
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-634-278-60

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWGGTNDYNTPF 14
Db 50 VIWGGTNDYNAAP 63

RESULT 5
US-07-634-278-61
; Sequence 61, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.

QY 1 VIWGGTNDYNTPF 14
Db 50 VIWGGTNDYNAAP 63

RESULT 3
US-07-634-278-37
; Sequence 37, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..119
; OTHER INFORMATION: /note="Amino acid sequence of the heavy chain of the humanized m1k-betal antibody."
; US-07-634-278-37

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWGGTNDYNTPF 14
Db 50 VIWGGTNDYNAAP 63

```

; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-634-278-61

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
Db 50 VIWSGGSTDYNAAF 63

RESULT 6
US-08-477-728-37
; Sequence 37, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,728
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-119
; OTHER INFORMATION: /note= "Amino acid sequence of the
; heavy chain of the humanized mik-beta1 antibody."
; US-08-477-728-37

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWSGGNTDYNTPF 14
Db 50 VIWSGGSTDYNAAF 63

RESULT 7
US-08-477-728-60
; Sequence 60, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US 07/310,252
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; CLASSIFICATION: 424
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990
;; APPLICATION DATA:
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/290,975
;; FILING DATE: 28-DEC-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 11823-002600
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 61:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 119 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-477-728-61

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPF 14
|||:|||||
Db 50 VIWGGSTDYNAF 63

RESULT 9
US-08-474-040-37
;; Sequence 37, Application US/08474040
;; Patent No. 5693761
;; GENERAL INFORMATION:
;; APPLICANT: QUEEN, Cary L.
;; APPLICANT: CO, Man Sung
;; APPLICANT: SCHNEIDER, William P.
;; APPLICANT: LANDOLFI, Nicholas F.
;; APPLICANT: COELING, Kathleen L.
;; APPLICANT: SELICK, Harold E.
;; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
;; NUMBER OF SEQUENCES: 113
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend Khourie and Crew
;; STREET: 379 Lytton Avenue
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: US
;; ZIP: 94301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,040
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/310,252
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/290,975
;; FILING DATE: 28-DEC-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,728
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990
;; APPLICATION DATA:
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/290,975
;; FILING DATE: 28-DEC-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 11823-002600
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 60:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 119 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-477-728-60

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VIWGGNTDYNTPF 14
|||:|||||
Db 50 VIWGGSTDYNAF 63

RESULT 8
US-08-477-728-61
;; Sequence 61, Application US/08477728
;; Patent No. 5585089
;; GENERAL INFORMATION:
;; APPLICANT: QUEEN, Cary L.
;; APPLICANT: SCHNEIDER, William P.
;; APPLICANT: SELICK, Harold E.
;; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
;; NUMBER OF SEQUENCES: 113
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: US
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,728
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990

```
/
/ REGISTRATION NUMBER: 30,223
/ REFERENCE/DOCKET NUMBER: 11823-002600
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 326-2400
/ TELEFAX: (415) 326-2422
/ INFORMATION FOR SEQ ID NO: 37:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 119 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..119
/ OTHER INFORMATION: /note= "Amino acid sequence of the
/ OTHER INFORMATION: heavy chain of the humanized mik-beta1 antibody."
/ US-08-474-040-37
/
Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VWSGGNTDYNTPF 14
Db 50 VWSGGSTDYNAAF 63

RESULT 10
US-08-474-040-60
; Sequence 60, Application US/08474040
; Patent No. 5693761
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO. Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,040
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; REFERENCE/DOCKET NUMBER: 11823-002600
```

```
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 326-2400
/ TELEFAX: (415) 326-2422
/ INFORMATION FOR SEQ ID NO: 60:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 119 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-474-040-60
/
Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VWSGGNTDYNTPF 14
Db 50 VWSGGSTDYNAAF 63

RESULT 11
US-08-474-040-61
; Sequence 61, Application US/08474040
; Patent No. 5693761
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO. Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,040
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; REFERENCE/DOCKET NUMBER: 11823-002600
```

OTHER INFORMATION: heavy chain of the humanized mik-betal antibody."
US-08-487-200-37
Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 VIWGGNTDYNTPF 14
DB 50 VIWGGSTDYNAAF 63
RESULT 12
US-08-487-200-37
Sequence 37, Application US/08487200
Patent No. 5693762
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..119
OTHER INFORMATION: /note= "Amino acid sequence of the

OTHER INFORMATION: heavy chain of the humanized mik-betal antibody."
US-08-487-200-37
Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 VIWGGNTDYNTPF 14
DB 50 VIWGGSTDYNAAF 63
RESULT 13
US-08-487-200-60
Sequence 60, Application US/08487200
Patent No. 5693762
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-487-200-60
Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 VWSGGNTDYNTPF 14
Db 50 VWSGGSTDYNAAF 63

RESULT 14
US-08-487-200-61
; Sequence 61, Application US/08487200
; Patent No. 5693762
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,200
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-487-200-61

Query Match 73.9%; Score 65; DB 1; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPF 14
Db 50 VWSGGSTDYNAAF 63

RESULT 15
US-08-484-537-37
; Sequence 37, Application US/08484537
; Patent No. 6180370
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,537
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: protein
; LOCATION: 1..119
; OTHER INFORMATION: /note="Amino acid sequence of the
; OTHER INFORMATION: heavy chain of the humanized mik-beta1 antibody."
; US-08-484-537-37

Query Match 73.9%; Score 65; DB 3; Length 119;
Best Local Similarity 78.6%; Pred. No. 0.018;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VWSGGNTDYNTPF 14
Db 50 VWSGGSTDYNAAF 63

Search completed: October 6, 2004, 16:36:31
Job time : 27:5789 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:59 ; Search time 11.5789 Seconds
(without alignments)
91.382 Million cell updates/sec

Title: US-09-635-974A-6
Perfect score: 62
Sequence: 1 LTYVDYEFAY 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 41 | 66.1 | 130 | C90363 | conserved hypother |
| 2 | 41 | 66.1 | 1020 | T42229 | probable E1-E2 ATP |
| 3 | 40 | 64.5 | 1009 | C64483 | hypothetical prote |
| 4 | 38 | 61.3 | 2204 | 1 RRMZNV | genome polyprotein |
| 5 | 37 | 59.7 | 121 | 2 S09558 | Ig heavy chain V-D |
| 6 | 37 | 59.7 | 130 | 2 A90230 | conserved hypother |
| 7 | 37 | 59.7 | 247 | 2 H71697 | hypothetical prote |
| 8 | 37 | 59.7 | 299 | 2 S35315 | cell division cont |
| 9 | 37 | 59.7 | 306 | 1 CEECDL | D-alanine-D-alanin |
| 10 | 37 | 59.7 | 306 | 2 H90640 | D-alanine-D-alanin |
| 11 | 37 | 59.7 | 306 | 2 H85491 | D-alanine-D-alanin |
| 12 | 37 | 59.7 | 306 | 2 AG0518 | D-alanine-D-alanin |
| 13 | 37 | 59.7 | 372 | 2 T24569 | hypothetical prote |
| 14 | 37 | 59.7 | 606 | 2 G71853 | flagellar hook-ass |
| 15 | 37 | 59.7 | 606 | 2 G64659 | flagellar hook-ass |
| 16 | 37 | 59.7 | 642 | 2 E72215 | oligopeptide ABC t |
| 17 | 36 | 58.1 | 62 | 2 C97865 | proline/betaine tr |
| 18 | 36 | 58.1 | 119 | 2 F27888 | Ig heavy chain v r |
| 19 | 36 | 58.1 | 130 | 2 F90278 | conserved hypother |
| 20 | 36 | 58.1 | 249 | 2 T17878 | hypothetical prote |
| 21 | 36 | 58.1 | 276 | 2 AC0248 | hypothetical prote |
| 22 | 36 | 58.1 | 335 | 2 A70107 | probable glycerald |
| 23 | 36 | 58.1 | 337 | 2 T32099 | hypothetical prote |
| 24 | 36 | 58.1 | 348 | 2 A70311 | hypothetical prote |
| 25 | 36 | 58.1 | 379 | 2 A70825 | probable 3-phenylp |
| 26 | 36 | 58.1 | 428 | 2 A83005 | conserved hypother |
| 27 | 36 | 58.1 | 659 | 2 A45184 | protein-tyrosine k |
| 28 | 36 | 58.1 | 659 | 2 I49553 | protein-tyrosine k |
| 29 | 36 | 58.1 | 1123 | 2 T28139 | PK4 protein kinase |

Query Match 66.1%; Score 41; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LTYVDYEFAY 11
Db 89 LTYVDASYAY 98

RESULT 2

T42229
Probable E1-E2 ATPase (EC 3.6.1.-) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C:Accession: T42229
R:Halleck, M.S.; Blackman, C.F.; Gao, L.; Williamson, P.L.; Schlegel, R.A.
submitted to the EMBL Data Library, June 1997
A:Description: Multiple members of a third subfamily of P-type ATPases identified by ge
A:Reference number: 222089
A:Accession: T42229
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1020 <HAL>
A:Cross-references: EMBL:AF011336; NID:G2944186; PID:g2944187; PIDN:AAC05245.1
C:Keywords: hydrolase; transmembrane protein

Query Match 66.1%; Score 41; DB 2; Length 1020;
Best Local Similarity 87.5%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LTYVDYEF 9
| | | | |

ALIGNMENTS

RESULT 1

C90363
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: C90363
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan,
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, I.
arrett, R.A.; Ragan, M.A.; Szensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: C90363
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-130 <KUR>
A:Cross-references: GB:AB006641; NID:g13815246; PIDN:AAK42162.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS01970

Query Match 66.1%; Score 41; DB 2; Length 130;
Best Local Similarity 70.0%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LTYVDYEFAY 11
| | | | | : | |
Db 89 LTYVDASYAY 98

30 36 58.1 1503 2 T43166
31 35.5 57.3 449 2 H69862
32 35 56.5 89 2 S24341
33 35 56.5 95 1 W5WL42
34 35 56.5 114 2 T16365
35 35 56.5 169 2 G71944
36 35 56.5 171 2 B4655
37 35 56.5 264 2 F64736
38 35 56.5 264 2 B90646
39 35 56.5 264 2 B85497
40 35 56.5 264 2 AH0524
41 35 56.5 301 2 G69961
42 35 56.5 301 2 S63620
43 35 56.5 347 2 T40599
44 35 56.5 364 1 JT0741
45 35 56.5 395 2 G97272

alpha-2-macroglobu
Nat-transporting A
hypothetical prote
E5 protein - human
hypothetical prote
hypothetical prote
hypothetical prote
3-methyl-2-oxobuta
hypothetical prote
hypothetical prote
3-methyl-2-oxobuta
phosphoenolpyruvat
cymI protein - Kle
hypothetical prote
GTP-binding protei
hypothetical prote

```

Db      711 LKYDYEF 718

RESULT 3
C64483
hypothetical protein MJ1468 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 10-Oct-1997
C:Accession: C64483
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
R.; Reich, C.L.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A54300; MUID:96337999; PMID:8688087
A:Accession: C64483
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1009 <BUL>
A:Cross-references: GB:U67587; GB:L77117; NID:gi15921103; PID:gi1592109; TIGR:MJ1468; PID:gi
C:Genetics:
A:Map position: FOR1437031-1440060

Query Match      64.5%; Score 40; DB 2; Length 1009;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 TYDYEF 9
Db      399 TYDIEY 405

RESULT 4
RRNZV
genome polyprotein - Newcastle disease virus (strain Beaudette C)
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: Newcastle disease virus
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: A26747
R:Yusoff, K.; Millar, N.S.; Chambers, P.; Emerson, P.T.
Nucleic Acids Res. 15, 3961-3976, 1987
A:Title: Nucleotide sequence analysis of the L gene of Newcastle disease virus: homolog
A:Reference number: A93665; MUID:87230982; PMID:3035486
A:Accession: A26747
A:Molecule type: mRNA
A:Residues: 1-2204 <YUS>
A:Cross-references: GB:X05399; NID:gi60937; PIDN:CAA28985.1; PID:gi60939
C:Genetics:
A:Gene: L
C:Superfamily: parainfluenza virus RNA-directed RNA polymerase
C:Keywords: ATP; nucleotidyltransferase

Query Match      61.3%; Score 38; DB 1; Length 2204;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 TYDYEFAY 11
Db      878 TYFDSEFSY 886

RESULT 5
S09958
Ig heavy chain V-D-J region (IE10) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 23-Jul-1999
C:Accession: S09958
R:Reininger, L.; Shibata, T.; Ozaki, S.; Shirai, T.; Jaton, J.C.; Izui, S.
Eur. J. Immunol. 20, 771-777, 1990
A:Title: Variable region sequences of pathogenic anti-mouse red blood cell autoantibodie
A:Reference number: S09955; MUID:90269328; PMID:2347362
A:Accession: S09958

```

```

A:Molecule type: mRNA
A:Residues: 1-121 <REI>
A:Cross-references: EMBL:X51846; NID:gi55246; PIDN:CAA36139.1; PID:gi930214
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-100/Domain: immunoglobulin homology <IMM>

Query Match      59.7%; Score 37; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 YDYEFAY 11
Db      103 XYGYSFAY 110

RESULT 6
A90230
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: A90230
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozeta, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A90230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-130 <XUR>
A:Cross-references: GB:AE006641; NID:gi3813978; PIDN:AAK41096.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO0798

Query Match      59.7%; Score 37; DB 2; Length 130;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 LTYDYEFAY 11
Db      89 LTYVDASYVY 98

RESULT 7
H71697
hypothetical protein RP401 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: H71697
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: H71697
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-247 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:gi3868717; PIDN:CAA14858.1; PID:gi3860958
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP401
C:Superfamily: Rickettsia prowazekii hypothetical protein RP401

Query Match      59.7%; Score 37; DB 2; Length 247;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 YDYEFAY 11
Db      33 YDYEFAY 39

```

```

Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64731
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-306 <BLAT>
A:Cross-references: GB:AB000118; GB:U00096; NID:gl786262; PIDN:AA073203.1; PID:gl786280.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ddbB; ddbI
A:Map position: 2 min
A>Note: gene is located in a large cluster of genes that are involved in cell division
C:Function:
A>Description: catalyzes ATP-driven formation of alanyl-D-alanine from 2 alanine molecules
A:Pathway: cell wall synthesis
A>Note: two D-alanine-D-alanine ligases in E. coli (and S. typhimurium) encoded by two
the two enzymes display remarkably similar catalytic efficiencies and substrate specificities
C:Superfamily: D-alanine-D-alanine ligase
C:Keywords: cell wall synthesis; dimer; ligase; magnesium
F:63-74/Region: D-alanine-D-alanine ligase motif 1
F:245-276/Region: D-alanine-D-alanine ligase motif 2

Query Match 59.7%; Score 37; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 10
H90640
D-alanine-D-alanine ligase B [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90640
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90640
A:Molecule type: DNA
A:Residues: 1-306 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033519.1; PID:gl3359552; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs0096
C:Superfamily: D-alanine-D-alanine ligase

Query Match 59.7%; Score 37; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 11
H85491
D-alanine-D-alanine ligase B ddbB [similarity] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Nov-2001
C:Accession: H85491
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551

```

```

Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64731
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-306 <BLAT>
A:Cross-references: GB:AB000118; GB:U00096; NID:gl786262; PIDN:AA073203.1; PID:gl786280.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ddbB; ddbI
A:Map position: 2 min
A>Note: gene is located in a large cluster of genes that are involved in cell division
C:Function:
A>Description: catalyzes ATP-driven formation of alanyl-D-alanine from 2 alanine molecules
A:Pathway: cell wall synthesis
A>Note: two D-alanine-D-alanine ligases in E. coli (and S. typhimurium) encoded by two
the two enzymes display remarkably similar catalytic efficiencies and substrate specificities
C:Superfamily: D-alanine-D-alanine ligase
C:Keywords: cell wall synthesis; dimer; ligase; magnesium
F:63-74/Region: D-alanine-D-alanine ligase motif 1
F:245-276/Region: D-alanine-D-alanine ligase motif 2

Query Match 59.7%; Score 37; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 10
H90640
D-alanine-D-alanine ligase B [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90640
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90640
A:Molecule type: DNA
A:Residues: 1-306 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033519.1; PID:gl3359552; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs0096
C:Superfamily: D-alanine-D-alanine ligase

Query Match 59.7%; Score 37; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 11
H85491
D-alanine-D-alanine ligase B ddbB [similarity] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Nov-2001
C:Accession: H85491
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551

```

```

Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64731
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-306 <BLAT>
A:Cross-references: GB:AB000118; GB:U00096; NID:gl786262; PIDN:AA073203.1; PID:gl786280.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ddbB; ddbI
A:Map position: 2 min
A>Note: gene is located in a large cluster of genes that are involved in cell division
C:Function:
A>Description: catalyzes ATP-driven formation of alanyl-D-alanine from 2 alanine molecules
A:Pathway: cell wall synthesis
A>Note: two D-alanine-D-alanine ligases in E. coli (and S. typhimurium) encoded by two
the two enzymes display remarkably similar catalytic efficiencies and substrate specificities
C:Superfamily: D-alanine-D-alanine ligase
C:Keywords: cell wall synthesis; dimer; ligase; magnesium
F:63-74/Region: D-alanine-D-alanine ligase motif 1
F:245-276/Region: D-alanine-D-alanine ligase motif 2

Query Match 59.7%; Score 37; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 10
H90640
D-alanine-D-alanine ligase B [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90640
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90640
A:Molecule type: DNA
A:Residues: 1-306 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033519.1; PID:gl3359552; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs0096
C:Superfamily: D-alanine-D-alanine ligase

Query Match 59.7%; Score 37; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
|:|||||
Db 208 TFYDYEAY 216

RESULT 11
H85491
D-alanine-D-alanine ligase B ddbB [similarity] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Nov-2001
C:Accession: H85491
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551

```

A;Accession: H85491
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-306 <STO>
 A;Cross-references: GB:AE005174; NID:G12512798; PIDN:AG054396.1; GSPDB:GN00145; UWGP:Z01
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: gdlB
 C;Superfamily: D-alanine-D-alanine ligase

Query Match 59.7%; Score 37; DB 2; Length 306;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TYVDYEFAY 11
 |:|||||
 Db 208 TFYDYEAKY 216

RESULT 12

AG0518
 D-alanine,D-alanine ligase B [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AG0518
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mouton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AG0518
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-306 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD01287.1; PID:G16501415; GSPDB:GN00176
 A;Gene: STY0150
 C;Superfamily: D-alanine-D-alanine ligase

Query Match 59.7%; Score 37; DB 2; Length 306;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TYVDYEFAY 11
 |:|||||
 Db 208 TFYDYEAKY 216

RESULT 13

T24569
 Hypothetical protein T06C12.4 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C;Accession: T24569
 R;Kelly, P.
 submitted to the EMBL Data Library, October 1996
 A;Reference number: Z1908
 A;Accession: T24569
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-372 <WIL>
 A;Cross-references: EMBL:Z81116; PIDN:CAB03304.1; GSPDB:GN00023; CESP:T06C12.4
 A;Experimental source: clone T06C12
 C;Genetics:
 A;Gene: CESP:T06C12.4
 A;Map position: 5
 A;Introns: 66/2; 290/3
 C;Superfamily: Caenorhabditis elegans hypothetical protein C36C9.3

Query Match 59.7%; Score 37; DB 2; Length 372;

Best Local Similarity 55.6%; Pred. No. 66;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALTYYDYEF 9
 |:|||||
 Db 41 AMTHYDFDF 49

RESULT 14

G71853
 flagellar hook-associated protein 1 (hap1) - Helicobacter pylori (strain J99)
 C;Species: Helicobacter pylori
 A;Variety: strain J99
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
 C;Accession: G71853
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, C.F.; Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
 A;Reference number: A71800; MUID:99120557; PMID:9923682
 A;Accession: G71853
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-606 <ARN>
 A;Cross-references: GB:AE001533; GB:AE001439; NID:G4155636; PIDN:AAD06640.1; PID:G4155655;
 A;Experimental source: strain J99
 C;Genetics:
 A;Gene: flgK

Query Match 59.7%; Score 37; DB 2; Length 606;
 Best Local Similarity 66.7%; Pred. No. 11e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TYVDYEFAY 11
 |:|||||
 Db 86 TYVDTEFSH 94

RESULT 15

G64659
 flagellar hook-associated protein 1 - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C;Accession: G64659
 R;Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Bordovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N.; A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A;Reference number: A64520; MUID:97394467; PMID:9252185
 A;Accession: G64659
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-606 <ROM>
 A;Cross-references: GB:AE000618; GB:AE000511; NID:G2314268; PIDN:AAD08163.1; PID:G2314277

Query Match 59.7%; Score 37; DB 2; Length 606;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TYVDYEFAY 11
 |:|||||
 Db 86 TYVDTEFSH 94

Search completed: October 6, 2004, 16:34:50
 Job time : 15:5789 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 6.75439 Seconds
(without alignments)
84,800 Million cell updates/sec

Title: US-09-635-974A-6

Perfect score: 62

Sequence: 1 ALTYDYEFAY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 43 | 69.4 | 358 | 1 CK33_CAVPO | Q92413 cavia porce |
| 2 | 40 | 64.5 | 1009 | 1 Y688_VETUA | Q58863 methanococ |
| 3 | 38 | 61.3 | 2204 | 1 RRPD_NDVB | P11205 newcastle d |
| 4 | 37 | 59.7 | 247 | 1 Y401_RICPR | Q924d0 rickettsia |
| 5 | 37 | 59.7 | 299 | 1 CC16_SCHPO | P36618 schizosacch |
| 6 | 37 | 59.7 | 305 | 1 DDLE_ECO57 | Q8X9Y6 escherichia |
| 7 | 37 | 59.7 | 305 | 1 DDLE_ECOL6 | Q8F163 escherichia |
| 8 | 37 | 59.7 | 305 | 1 DDLE_ECOLI | P07862 escherichia |
| 9 | 37 | 59.7 | 305 | 1 DDLE_SALTI | Q82997 salmonella |
| 10 | 37 | 59.7 | 305 | 1 DDLE_SALTY | Q82911 salmonella |
| 11 | 37 | 59.7 | 306 | 1 DDLE_SHIFL | Q83mf7 shigella fl |
| 12 | 37 | 59.7 | 1047 | 1 AT9A_HUMAN | O75110 homo sapien |
| 13 | 37 | 59.7 | 1047 | 1 AT9A_MOUSE | O70228 mus musculu |
| 14 | 36 | 58.1 | 335 | 1 G3P_BORBU | P46795 borrelia bu |
| 15 | 36 | 58.1 | 348 | 1 Y113_AQUAE | O68516 aquifex aeo |
| 16 | 36 | 58.1 | 416 | 1 HGB_LGPN | Q984t0 legionella |
| 17 | 36 | 58.1 | 659 | 1 BTK_HUMAN | Q06187 homo sapien |
| 18 | 36 | 58.1 | 659 | 1 BTK_MOUSE | P35991 mus musculu |
| 19 | 36 | 58.1 | 724 | 1 FCRA_BACST | P56255 bacillus st |
| 20 | 35 | 56.5 | 95 | 1 VES_HV42 | P27227 human papil |
| 21 | 35 | 56.5 | 160 | 1 COAD_LEPIN | Q82YP6 leptospira |
| 22 | 35 | 56.5 | 263 | 1 PANB_SALTY | Q82910 salmonella |
| 23 | 35 | 56.5 | 264 | 1 PANB_ECO57 | Q8X929 escherichia |
| 24 | 35 | 56.5 | 264 | 1 PANB_ECOL6 | Q8F130 escherichia |
| 25 | 35 | 56.5 | 264 | 1 PANB_ECOLI | P31057 escherichia |
| 26 | 35 | 56.5 | 264 | 1 PANB_SALTI | Q829d2 salmonella |
| 27 | 35 | 56.5 | 301 | 1 CPM_BACSU | P54528 bacillus su |
| 28 | 35 | 56.5 | 303 | 1 DDLE_CHRVO | Q7Aq01 chromobacte |
| 29 | 35 | 56.5 | 308 | 1 Y308_MYCGA | P33275 mycoplasma |
| 30 | 35 | 56.5 | 338 | 1 LUM_HUMAN | P51884 homo sapien |
| 31 | 35 | 56.5 | 364 | 1 GTF1_SCHPO | P32235 schizosacch |
| 32 | 35 | 56.5 | 421 | 1 PTE2_HUMAN | P49753 homo sapien |
| 33 | 35 | 56.5 | 535 | 1 VAN1_YEAST | P23642 saccharomyc |

34 35 56.5 881 1 HELI_VZVD
35 35 56.5 917 1 SLAP_THETH
36 34 54.8 124 1 PYP_RHOSH
37 34 54.8 144 1 YG34_BPLLH
38 34 54.8 144 1 YG34_BPMV4
39 34 54.8 179 1 Y256_SULSO
40 34 54.8 226 1 YEDJ_LACLA
41 34 54.8 304 1 DDL_NEIMA
42 34 54.8 304 1 DDL_NEIMB
43 34 54.8 312 1 G3P_MESAU
44 34 54.8 319 1 DDLB_PSEAE
45 34 54.8 320 1 G3P_BOVIN

ALIGNMENTS

RESULT 1
CKR3_CAVPO
ID CKR3_CAVPO STANDARD; PRT; 358 AA.
AC Q922i3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE C-C chemokine receptor type 3 (C-C CKR-3) (CC-CKR-3) (CCR3)
DE (CCR3)
GN CCR3 OR CMKR3.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95049845; PubMed=9834099;
RA Sabroe I., Conroy D.M., Gerard N.P., Li Y., Collins P.D., Post T.W.,
RA Jose P.J., Williams T.J., Gerard C.J., Ponath P.D.;
RT "Cloning and characterization of the guinea pig eosinophil eotaxin
RT receptor, C-C chemokine receptor-3: blockade using a monoclonal
RT antibody in vivo."
RL J. Immunol. 161:6139-6147(1998).
CC -!- FUNCTION: Receptor for a C-C type chemokine. Binds to eotaxin,
CC MCP-3, MCP-4 and RANTES and subsequently transduces a signal by
CC increasing the intracellular calcium ions level (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AF060698; AAC80428.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 44 64 POTENTIAL.
FT DOMAIN 65 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 95 POTENTIAL.
FT DOMAIN 96 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 133 POTENTIAL.
FT DOMAIN 134 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 POTENTIAL.
FT DOMAIN 176 206 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 207 227 POTENTIAL.
FT DOMAIN 228 243 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 244 264 POTENTIAL.

FT DOMAIN 265 287 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 288 308 POTENTIAL.
 FT DOMAIN 309 358 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 358 AA; 41623 MW; 7B73FAB7A3BC3670 CRC64;

Query Match 69.4%; Score 43; DB 1; Length 358;
 Best Local Similarity 87.5%; Pred. No. 3.4;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 TYDYDEFA 10
 DB 18 TYDYDEFA 25

RESULT 2
 ID Y668 METJA STANDARD; PRT; 1009 AA.

AC Q5863; (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein MJ1468.

GN MJ1468.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kleravage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Contains 5 PKD domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U67588; AAB39478.1; -
 CC PIR; C64483; C64483.
 CC TIGR; MJ1468; -

DR InterPro: IPR000601; PKD.
 DR Pfam; PF00801; PKD; 5.
 DR SMART; SM00089; PKD; 7.
 DR PROSITE; PS50093; PKD; 5.

KW Hypothetical protein; Transmembrane; Repeat; Complete proteome.

FT TRANSMEM 6 26 POTENTIAL.
 FT TRANSMEM 985 1005 POTENTIAL.
 FT DOMAIN 213 247 PKD 1.
 FT DOMAIN 436 503 PKD 2.
 FT DOMAIN 724 806 PKD 3.
 FT DOMAIN 822 886 PKD 4.
 FT DOMAIN 925 962 PKD 5.
 FT DOMAIN 293 298 POLY-ASN.
 SQ SEQUENCE 1009 AA; 115119 MW; 13E9B4933EAB7972 CRC64;

Query Match 64.5%; Score 40; DB 1; Length 1009;
 Best Local Similarity 85.7%; Pred. No. 31;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TYDYDEF 9
 DB 399 TYDYDEF 405

RESULT 3
 RRPL NDVB STANDARD; PRT; 2204 AA.

AC P11255;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
 DE (L protein).
 GN L.
 OS Newcastle disease virus (strain Beaudette C/45) (NDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
 OX NCBI_TaxID=11178;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=87230982; PubMed=3035486;
 RA Yusoff K., Millar N.S., Chambers P., Emmerson P.T.;
 RT "Nucleotide sequence analysis of the L gene of Newcastle disease
 RT virus: homologies with Sendai and vesicular stomatitis viruses.";
 CC Nucleic Acids Res. 15:3961-3976(1987).
 CC -1- FUNCTION: Probable component of the active polymerase. It may
 CC function in mRNA synthesis, capping, methylation and poly(A)
 CC synthesis of newly synthesized viral mRNAs, RNA editing of the P
 CC gene transcript, and protein kinase activity.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).

CC -1- SIMILARITY: Belongs to the paramyxoviruses L protein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X05399; CAA28985.1; -
 CC PIR; A26747; RSNZNV.
 CC InterPro: IPR007098; RNA_pol_monon.
 CC InterPro: IPR001016; Viral_RNA_pol_L.
 CC Pfam; PF00946; Paramyx_RNA_pol; 1.
 CC Transferase; RNA-directed RNA polymerase.

KW Transferase; RNA-directed RNA polymerase.
 SQ SEQUENCE 2204 AA; 248822 MW; C67B8674D904802C CRC64;

Query Match 61.3%; Score 38; DB 1; Length 2204;
 Best Local Similarity 56.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TYDYDEFAY 11
 DB 878 TYDSEFSY 886

RESULT 4
 Y401 RICPR STANDARD; PRT; 247 AA.

ID Y401 RICPR
 AC Q9ZDD0;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein RP401.
 GN RP401.

OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.

```

OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O., Naeslund A.K.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Kurland C.G.;
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ235271; CAA14858.1; -.
CC DR PIR; H71697; H71697.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 247 AA; 28125 MW; A30702F1912543D5 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 247;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 YDYEFAH 11
Db 33 YDYEFAH 39
|||||:

RESULT 5
CC16_SCHPO STANDARD; PRT; 299 AA.
ID CC16_SCHPO
AC P36618;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division control protein 16.
GN CDC16 OR SPAC6F6.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=93327759; PubMed=8334988;
RA Fankhauser C., Marks J., Raymond A., Simanis V.;
RT "The S. pombe cdc16 gene is required both for maintenance of p34cdc2
RT kinase activity and regulation of septum formation: a link between
RT mitosis and cytokinesis?";
RL EMBO J. 12:2697-2704(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins M., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

```

```

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Calbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -----
CC !- FUNCTION: HAS A DUAL ROLE IN THE CELL CYCLE. IN MITOSIS, IT IS
CC INVOLVED IN MAINTENANCE OF CDC2 KINASE ACTIVITY. IT IS
CC SUBSEQUENTLY REQUIRED FOR REGULATION OF SEPTUM FORMATION. COULD BE
CC INVOLVED IN MAINTENANCE OF CDC2 KINASE ACTIVITY BY PREVENTING,
CC DIRECTLY OR INDIRECTLY, THE DEGRADATION OF CYCLIN OR THE
CC DEPHOSPHORYLATION OF THR-167 OF CDC2.
CC !- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X71605; CAA50606.1; -.
CC DR EMBL; Z98981; CAB11731.1; -.
CC DR PIR; S35315; S35315.
CC DR GeneDB_Spomb; SPAC6F6.08C; -.
CC DR InterPro; IPR000195; RabGAP_TBC.
CC DR Pfam; PF00566; TBC; 1.
CC DR SMART; SM00164; TBC; 1.
CC DR PROSITE; PS50086; TBC_RABGAP; 1.
CC KW Cell cycle.
CC FT DOMAIN 41 223 RAB-GAP TBC.
CC SQ SEQUENCE 299 AA; 33988 MW; 8959846E4032766D CRC64;

Query Match 59.7%; Score 37; DB 1; Length 299;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
Db 212 ALTYDYEFAY 222
|||||:

RESULT 6
DDLB_ECO57 STANDARD; PRT; 305 AA.
ID DDLB_ECO57
AC Q8X9Y6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-CCT-2003 (Rel. 42, Last annotation update)
DE D-alanine-D-alanine ligase B (EC 6.3.2.4) (D-alanylalanine
DE synthetase B) (D-Ala-D-Ala ligase B).
GN DDLB OR J0102 OR ECS0096.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,

```

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7."
 RN Nature 409:529-533 (2001).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Kurokawa K., Murata T., Tanaka M., Tobe T.,
 RA Han C.-G., Ohsuho E., Nakayama K.,
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: Cell wall formation (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
 CC alanyl-D-alanine.
 CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
 CC step.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: A005186; AAG54396.1; -;
 CC EMBL: A005250; BAB33519.1; -;
 CC PIR: H85491; H85491.
 CC PIR: H90640; H90640.
 CC HAMAP: MF_00047; -; 1.
 CC InterPro: IPR005905; D_alan_ligase.
 CC InterPro: IPR000291; Dala_lig_Van.
 CC Pfam: PF01820; Dala_ligase; 1.
 CC TIGRFAMs: TIGR01205; D_alan_ligase; 1.
 CC PROSITE: PS00843; DALA_DALA_LIGASE 1; 1.
 CC PROSITE: PS00844; DALA_DALA_LIGASE 2; 1.
 CC Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
 KW INIT MET 0 BY SIMILARITY.
 FT ACT_SITE 14 14 BY SIMILARITY.
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT ACT_SITE 280 280 BY SIMILARITY.
 SQ SEQUENCE 305 AA; 32722 MW; B8C61308C79F36F1 CRC64;
 Query Match 59.7%; Score 37; DB 1; Length 305;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 TYVDYEFAY 11
 Db 207 TFYDYEAKY 215
 RESULT 7
 ID DDLB_ECOL6 STANDARD; PRT; 305 AA.
 AC Q8FL63;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE D-alanine--D-alanine ligase B (EC 6.3.2.4) (D-alanylalanine
 synthetase B) (D-Ala-D-Ala ligase B).
 GN DDLB OR C0110.
 OS *Escherichia coli* O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxID=217992;

RN SEQUENCE FROM N.A.
 RP STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=2238234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.,
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 CC -!- FUNCTION: Cell wall formation (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
 CC alanyl-D-alanine.
 CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
 CC step.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AE016755; AAN78608.1; -;
 CC HAMAP: MF_00047; -; 1.
 CC InterPro: IPR005905; D_alan_ligase.
 CC InterPro: IPR000291; Dala_lig_Van.
 CC Pfam: PF01820; Dala_ligase; 1.
 CC TIGRFAMs: TIGR01205; D_alan_ligase; 1.
 CC PROSITE: PS00843; DALA_DALA_LIGASE 1; 1.
 CC PROSITE: PS00844; DALA_DALA_LIGASE 2; 1.
 CC Ligase; Cell wall; Peptidoglycan synthesis; Complete proteome.
 KW INIT MET 0 BY SIMILARITY.
 FT ACT_SITE 14 14 BY SIMILARITY.
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT ACT_SITE 280 280 BY SIMILARITY.
 SQ SEQUENCE 305 AA; 32761 MW; E09D9604F7D5BF0F CRC64;
 Query Match 59.7%; Score 37; DB 1; Length 305;
 Best Local Similarity 66.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 3 TYVDYEFAY 11
 Db 207 TFYDYEAKY 215
 RESULT 8
 ID DDLB_ECOLI STANDARD; PRT; 305 AA.
 AC P07862;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE D-alanine--D-alanine ligase B (EC 6.3.2.4) (D-alanylalanine
 synthetase B) (D-Ala-D-Ala ligase B).
 GN DDLB OR DDL OR B0092.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxID=562;
 RN [1].
 RP SEQUENCE FROM N.A.
 RX STRAIN=K12;
 RX MEDLINE=86304170; PubMed=3528126;
 RA Robinson A.C., Kenan D.J., Sweeney J., Donachie W.D.,
 RT "Further evidence for overlapping transcriptional units in an
Escherichia coli cell envelope-cell division gene cluster: DNA

sequence and transcriptional organization of the ddl ftsQ region.";
J. Bacteriol. 167:809-817(1986).
[2]
SEQUENCE FROM N.A.
STRAIN=K12;
MEDLINE=92334977; PubMed=1630901;
Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
Isono K., Mizobuchi K., Nakata A.;
"Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.";
Nucleic Acids Res. 20:3305-3308(1992).
[3]
SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Ferna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
"The complete genome sequence of Escherichia coli K-12.";
Science 277:1232-1244(1997).
[4]
SEQUENCE OF 1-40 FROM N.A.
STRAIN=K12;
MEDLINE=90326550; PubMed=2197603;
Ikeda M., Wachi M., Jung H.K., Ishino F., Matsuhashi M.;
"Nucleotide sequence involving murG and murC in the mra gene cluster
region of Escherichia coli.";
Nucleic Acids Res. 18:4014-4014(1990).
[5]
CHARACTERIZATION, AND PARTIAL SEQUENCE.
MEDLINE=92207163; PubMed=1554356;
Al-Bar O.A., O'Connor C.D., Giles I.G., Akhtar M.;
"D-alanine:D-alanine ligase of Escherichia coli. Expression,
purification and inhibitory studies on the cloned enzyme.";
Biochem. J. 282:747-752(1992).
[6]
X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
MEDLINE=95025939; PubMed=7939684;
Pan C., Moews P.C., Walsh C.T., Knox J.R.;
"Vancomycin resistance: structure of D-alanine:D-alanine ligase at
2.3-A resolution.";
Science 266:439-443(1994).
[7]
X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
MEDLINE=97207065; PubMed=9054558;
Pan C., Park I.-S., Walsh C.T., Knox J.R.;
"D-alanine:D-alanine ligase: phosphonate and phosphinate
intermediates with wild type and the Y216F mutant.";
Biochemistry 36:2531-2538(1997).
CC -!- FUNCTION: Cell wall formation.
CC -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
CC -!- alanyl-D-alanine.
CC -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC step.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the D-alanine--D-alanine ligase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to licens@isb-sib.ch).

EMBL; M14029; AAA23672.1; -
EMBL; X02668; AAA23815.1; -
EMBL; X52644; CAA33869.1; -
EMBL; X55034; CAA38869.1; -
EMBL; D10483; BAB96660.1; -
EMBL; AE000118; AAC73203.1; -.

DR PIR; A30289; CEECDL.
DR PDB; 2DLN; 01-NOV-95.
DR PDB; 11OV; 12-FEB-97.
DR PDB; 11OW; 12-FEB-97.
DR EcoGene; EG10214; ddb.
DR HAMAP; MF_00047; -; 1.
DR InterPro; IPR005905; Dala D ala.
DR InterPro; IPR000291; Dala Lig Van.
DR Pfam; PF01820; Dala Dala Ligas; 1.
DR TIGRPFAMs; TIGR01205; D ala D ala IGR; 1.
DR PROSITE; PS00843; DALA DALA LIGASE_1; 1.
DR PROSITE; PS00844; DALA DALA LIGASE_2; 1.
KW Ligase; Cell wall; Peptidoglycan synthesis; 3D-structure;
KW Complete proteome.
FT INIT MET 0
FT ACT_SITE 14
FT ACT_SITE 149
FT ACT_SITE 280
FT STRAND 3
FT STRAND 7
FT TURN 13
FT HELIX 15
FT TURN 31
FT TURN 32
FT STRAND 35
FT TURN 41
FT HELIX 46
FT TURN 49
FT STRAND 54
FT TURN 64
FT HELIX 70
FT TURN 79
FT STRAND 82
FT HELIX 87
FT HELIX 96
FT TURN 106
FT STRAND 110
FT STRAND 111
FT HELIX 118
FT TURN 127
FT TURN 127
FT HELIX 128
FT TURN 135
FT TURN 140
FT STRAND 140
FT TURN 145
FT TURN 149
FT STRAND 154
FT HELIX 159
FT HELIX 162
FT TURN 170
FT TURN 172
FT STRAND 175
FT STRAND 186
FT TURN 193
FT STRAND 195
FT STRAND 200
FT HELIX 211
FT TURN 216
FT STRAND 221
FT HELIX 230
FT TURN 248
FT STRAND 252
FT TURN 261
FT STRAND 265
FT TURN 278
FT HELIX 281
FT TURN 289
FT HELIX 293
FT TURN 303
SQ SEQUENCE 305 AA; 32708 MW; 79103A85E732A4C7 CRC64;
Query Match 59.7%; Score 37; DB 1; Length 305;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 TTYDYEFAY 11
|:|||||


```

DDL8_SHIFL          STANDARD;          PRT;   306 AA.
ID  DBL8_SHIFL      Q7UDS4;
AC  Q83MF7; Q7UDS4;
DT  15-MAR-2004 (Rel. 43, Created)
DT  15-MAR-2004 (Rel. 43, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  D-alanine-D-alanine ligase B (EC 6.3.2.4) (D-alanylanine
DE  synthetase B) (D-Ala-D-Ala ligase B).
GN  DDL8 OR SF0089 OR S0091.
OS  Shigella flexneri.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Shigella.
OX  NCBI_TaxID=623;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=301 / Serotype 2a;
RX  MEDLINE=22272406; PubMed=12384590;
RA  Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA  Yang J., Goldberg M.B., Burland V., Zhang J., Yang G., Wu H., Qi D., Dong J.,
RA  Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA  Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA  Yu J.;
RT  "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT  through comparison with genomes of Escherichia coli K12 and O157."
RL  Nucleic Acids Res. 30:4432-4441(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX  MEDLINE=42590274; PubMed=12704152;
RA  Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA  Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA  Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA  Schwartz D.C., Blattner F.R.;
RT  "Complete genome sequence and comparative genomics of Shigella
RT  flexneri serotype 2a strain 2457T."
RL  Infect. Immun. 71:2775-2796(2003).
RN  [1]
RP  FUNCTION: Cell wall formation (By similarity).
CC  -!- CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-
CC  alanyl-D-alanine.
CC  -!- PATHWAY: D-alanine branch of peptidoglycan biosynthesis; second
CC  step.
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -!- SIMILARITY: Belongs to the D-alanine-D-alanine ligase family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AE015047; AAN41754.1; -
CC  EMBL; AE016978; AAP15635.1; -
CC  HAMAP; MF_00047; 1.
CC  InterPro; IPR000291; Dala_lig_Van.
CC  InterPro; IPR005905; D_alal_ligase.
CC  Pfam; PF01820; Dala_dala_ligase; 1.
CC  TIGRPFAMs; TIGR01205; D_alal_ligase; 1.
CC  PROSITE; PS00843; DALA_DALA_LIGASE_1; 1.
CC  PROSITE; PS00844; DALA_DALA_LIGASE_2; 1.
CC  LIGASE; Cell wall; Peptidoglycan synthesis; Complete proteome.
CC  CONFLICT 65 65 R -> P (IN REF. 2).
CC  SEQUENCE 306 AA; 32821 MW; A07A81804D4AADB6 CRC64;
Query Match          59.7%; Score 37; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 TYDYEFAY 11
Db 208 TYDYEAKY 216

```

RESULT 12

```

AT9A_HUMAN
ID  AT9A_HUMAN      STANDARD;          PRT;   1047 AA.
AC  Q751T0; Q9NQK6; Q9NQK7;
DT  30-MAY-2000 (Rel. 39, Created)
DT  16-MAY-2001 (Rel. 40, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Potential phospholipid-transporting ATPase IIA (EC 3.6.1.-).
GN  ATP9A OR KIAA0611.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE OF 15-1047 FROM N.A. (ISOFORM SHORT).
RC  TISSUE=Brain;
RX  MEDLINE=98403880; PubMed=9734811;
RA  Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA  Kotani H., Nomura N., Ohara O.;
RT  "Prediction of the coding sequences of unidentified human genes. X.
RT  The complete sequences of 100 new cDNA clones from brain which can
RT  code for large proteins in vitro."
RL  DNA Res. 5:169-176(1998).
RN  [2]
RP  SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RX  MEDLINE=21638749; PubMed=11780052;
RA  Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA  Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA  Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA  Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA  Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA  Chapman J.C., Clamp M., Clark L.N., Clark S.Y., Cooby N.R.,
RA  Clegg S., Cobley V.E., Collier R.E., Connor R.E., Dhami P.D., Dunn M.,
RA  Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA  Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA  Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA  Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA  Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA  Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA  Leharvalho M.H., Levensha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA  Marsh V.L., Martin S.L., McConnachie L.J., McIay K., McMurray A.A.,
RA  Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA  Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA  Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA  Rice C.M., Ross M.T., Scott C.E., Senra H.K., Showkeen R., Sims S.,
RA  Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA  Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA  Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA  Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA  Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA  Rogers J.;
RT  "The DNA sequence and comparative analysis of human chromosome 20."
RL  Nature 414:865-871(2001).
CC  -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC  -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC  -!- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=Long;
CC  IsoId=Q75110-1; Sequence=Displayed;
CC  Name=Short;
CC  IsoId=Q75110-2; Sequence=VSP_000432;
CC  -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC  ATPases). Subfamily IV.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

DR EMBL; AB014511; BAA31586.1; -
 DR EMBL; AL353799; CAC17554.1; -
 DR EMBL; AL138807; CAB92773.1; -
 DR EMBL; AL038684; CAB63450.1; -
 DR EMBL; AL138807; CAB92774.1; ALT_SEQ.
 DR Genew; HGNC:13540; ATP9A.
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR InterPro; IPR006539; Flippase.
 DR InterPro; IPR005834; Hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR TIGRFAMS; TIGR01652; ATPase-Plipid; 1.
 DR TIGRFAMS; TIGR01494; ATPase_P-type; 7.
 DR PROSITE; PS00154; ATPase_E1_E2; 1.
 DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KW Alternative splicing; Multigene family.
 FT DOMAIN 1 69
 FT TRANSMEM 70 91
 FT DOMAIN 92 96
 FT TRANSMEM 97 119
 FT DOMAIN 120 303
 FT TRANSMEM 304 325
 FT DOMAIN 326 332
 FT TRANSMEM 333 354
 FT DOMAIN 355 841
 FT TRANSMEM 842 862
 FT DOMAIN 863 874
 FT TRANSMEM 875 893
 FT DOMAIN 894 923
 FT TRANSMEM 924 942
 FT DOMAIN 943 949
 FT TRANSMEM 950 972
 FT DOMAIN 973 978
 FT TRANSMEM 979 999
 FT DOMAIN 1000 1006
 FT TRANSMEM 1007 1030
 FT DOMAIN 1031 1047
 FT MOD_RES 391 391
 FT METAL 785 785
 FT METAL 789 789
 FT VARSPLIC 149 269
 SQ SEQUENCE 1047 AA; 118582 MW; 80C307CF5A396755 CRC64;
 Query Match 59.7%; Score 37; DB 1; Length 1047;
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LTYDYEF 9
 Db 738 LKYYEYEF 745
 RESULT 13
 AT9A_MOUSE STANDARD; PRT; 1047 AA.
 ID AT9A_MOUSE Q922L9;
 AC Q922L9; Q922L9;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Potential phospholipid-transporting ATPase IIA (EC 3.6.3.1).
 OS ATP9A.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20473714; PubMed=11015572;
 RA Halleck M.S., Lawler J.F. Jr., Blackshaw S., Gao L., Nagarajan P.,
 RA Hacker C., Pyle S., Newman J.T., Nakanishi Y., Ando H., Weinstock D.,
 RA Williamson P., Schlegel R.A.;

RT "Differential expression of putative transbilayer amphipath
 RT transporters.";
 RL Physiol. Genomics 1:139-150 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Males J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Grimwood J., Schmutz J., Myers R.M.,
 RA Rodriguez A.C., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Butterfield Y.S.N., Jones S.J.M., Marra M.A.;
 RA Schlegel R.A.;
 RT "Multiple members of a third subfamily of P-type ATPases identified by
 RT genomic sequences and ESTs";
 RL Genome Res. 8:354-361 (1998).
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: Found in most tissues except spleen. Most
 CC abundant in brain. Also detected in fetal tissues.
 CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 CC ATPases). Subfamily IV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF152243; AAF08396.1; -
 DR EMBL; AF011336; AAC05245.1; -
 DR EMBL; BC021814; AAH21814.1; -
 DR EMBL; BC006949; AAH06949.1; -
 DR PIR; T42229; T42229.
 DR MGD; MGI:1330826; Atp9a.
 DR InterPro; IPR001757; ATPase_E1-E2.
 DR InterPro; IPR006539; Flippase.
 DR InterPro; IPR005834; Hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR TIGRFAMS; TIGR01652; ATPase-Plipid; 1.
 DR TIGRFAMS; TIGR01494; ATPase_P-type; 7.
 DR PROSITE; PS00154; ATPase_E1_E2; 1.
 DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KW Multigene family.
 FT DOMAIN 1 69
 FT TRANSMEM 70 91
 FT DOMAIN 92 96
 FT TRANSMEM 97 119
 FT DOMAIN 120 303
 FT TRANSMEM 304 325
 FT DOMAIN 326 332
 FT TRANSMEM 333 354
 FT TRANSMEM 333 354

FT DOMAIN 355 841 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 842 862 POTENTIAL.
FT DOMAIN 863 874 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 875 893 POTENTIAL.
FT DOMAIN 894 923 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 924 942 POTENTIAL.
FT DOMAIN 943 949 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 950 972 POTENTIAL.
FT DOMAIN 973 978 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 979 999 POTENTIAL.
FT DOMAIN 1000 1006 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1007 1030 POTENTIAL.
FT DOMAIN 1031 1047 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 391 391 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 785 785 MAGNESIUM (BY SIMILARITY).
FT METAL 789 789 MAGNESIUM (BY SIMILARITY).
FT CONFLICT 29 29 R -> K (IN REF. 3).
FT CONFLICT 129 129 V -> I (IN REF. 3).
FT CONFLICT 138 138 V -> I (IN REF. 3).
FT CONFLICT 227 227 K -> E (IN REF. 2).
FT CONFLICT 243 243 N -> D (IN REF. 2).
FT CONFLICT 278 278 R -> D (IN REF. 3).
FT CONFLICT 279 279 K -> E (IN REF. 2).
FT CONFLICT 432 432 Q -> P (IN REF. 3).
FT CONFLICT 475 475 S -> C (IN REF. 3).
FT CONFLICT 504 510 ALVOWTE -> TLAKREG (IN REF. 2; AAH21814).
FT CONFLICT 533 533 V -> I (IN REF. 3).
FT CONFLICT 613 613 H -> D (IN REF. 3).
FT CONFLICT 742 742 E -> D (IN REF. 3).
FT CONFLICT 784 784 W -> G (IN REF. 3).
FT SEQUENCE 1047 AA; 118756 MW; B76AEA98C32F6D04 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 1047;
Best Local Similarity 75.0%; Pred. No. 1.e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTYDYDEF 9
|:|:|:|
Db 738 LKYYEYEF 745

RESULT 14
G3P BORBU STANDARD; PRT; 335 AA.
AC P46795; O51084;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR B80057.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=8557349;
RA Anda P., Gebbia J.A., Backenson P.B., Coleman J.L., Benach J.L.;
"A glyceraldehyde-3-phosphate dehydrogenase homolog in Borrelia
burgdorferi and Borrelia hermsili";
RL Infect. Immun. 64:262-268 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Cajigas S., Huang W.M., Sutton G.G., Clayton P.A.,
Lathigra R., White O., Kechum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kellavag A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;

RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";
RL Nature 390:580-586 (1997).
CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -|- PATHWAY: Second phase of glycolysis; first step.
CC -|- SUBUNIT: Homotetramer (by similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; U28760; AB53930.1; -.
DR EMBL; AE001119; AAC66450.1; -.
DR PIR; A70107; A70107.
DR HSSP; P00362; 1GDI.
CC
CC TIGR; B80057; -.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; Gpdh; 1.
DR Pfam; PF02800; Gpdh; C; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 133 153 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT CONFLICT 214 214 A -> P (IN REF. 1).
FT CONFLICT 294 294 S -> P (IN REF. 1).
FT SEQUENCE 335 AA; 36254 MW; 30E94F98939819C0 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 335;
Best Local Similarity 60.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYDEF 11
|:|:|:|
Db 313 LSWYDNERGY 322

RESULT 15
Y113 AQUAE STANDARD; PRT; 348 AA.
AC O66516;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_113 precursor.
GN AQ_113.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus";
RL Nature 392:353-358 (1998).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE000674; AAC06480.1; -
 DR PIR; A70311; A70311.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 17
 FT CHAIN 18 348
 SQ SEQUENCE 348 AA; 38760 MW; BDDC09E0013AF152 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 348;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDY 7
 Db 208 ALSYDY 214

Search completed: October 6, 2004, 16:30:25
 Job time : 7.75439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:13 ; Search time 36.8596 Seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-6
Perfect score: 62
Sequence: 1 ALTYDYEFAY 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mnc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 44 | 71.0 | 2054 | Q813Q3 | Q813Q3 plasmodium |
| 2 | 43 | 69.4 | 259 | Q8P041 | Q8P041 streptococc |
| 3 | 43 | 69.4 | 259 | Q8K6N0 | Q8K6N0 streptococc |
| 4 | 41 | 66.1 | 130 | Q97WY9 | Q97WY9 sulfolobus |
| 5 | 41 | 66.1 | 210 | Q8E3E1 | Q8E3E1 streptococc |
| 6 | 41 | 66.1 | 210 | Q8DXS1 | Q8DXS1 streptococc |
| 7 | 40 | 64.5 | 323 | Q860W8 | Q860W8 mus musculu |
| 8 | 40 | 64.5 | 329 | Q85ZW9 | Q85ZW9 mus musculu |
| 9 | 39 | 62.9 | 77 | Q857E5 | Q857E5 mycobacteri |
| 10 | 39 | 62.9 | 495 | Q9CV82 | Q9CV82 mus musculu |
| 11 | 39 | 62.9 | 580 | Q89YL8 | Q89YL8 bacteroides |
| 12 | 38 | 61.3 | 148 | Q8EMU7 | Q8EMU7 ansacta moo |
| 13 | 38 | 61.3 | 154 | Q8XLE4 | Q8XLE4 clostridium |
| 14 | 38 | 61.3 | 255 | Q8P7X5 | Q8P7X5 xanthomonas |
| 15 | 38 | 61.3 | 333 | Q8JGW5 | Q8JGW5 rana ridibu |
| 16 | 38 | 61.3 | 465 | Q8SQJ3 | Q8SQJ3 encephalito |

| | | | | | | |
|----|----|------|------|----|--------|--------------------|
| 17 | 38 | 61.3 | 503 | 10 | Q93WC0 | Q93WC0 hordeum vul |
| 18 | 38 | 61.3 | 547 | 16 | Q89QL9 | Q89QL9 bradyrhizob |
| 19 | 38 | 61.3 | 721 | 16 | Q87J71 | Q87J71 vibrio para |
| 20 | 38 | 61.3 | 816 | 4 | Q8NB73 | Q8NB73 homo sapien |
| 21 | 38 | 61.3 | 1096 | 4 | Q94836 | Q94836 homo sapien |
| 22 | 38 | 61.3 | 1543 | 12 | Q09438 | Q09438 garlic viru |
| 23 | 38 | 61.3 | 1544 | 12 | Q8QX11 | Q8QX11 garlic viru |
| 24 | 38 | 61.3 | 1550 | 12 | Q55601 | Q55601 human immun |
| 25 | 37 | 59.7 | 22 | 15 | Q9PXA8 | Q9PXA8 human immun |
| 26 | 37 | 59.7 | 24 | 15 | Q9PXA7 | Q9PXA7 human immun |
| 27 | 37 | 59.7 | 26 | 15 | Q9PXA6 | Q9PXA6 human immun |
| 28 | 37 | 59.7 | 130 | 17 | Q9UX14 | Q9UX14 sulfolobus |
| 29 | 37 | 59.7 | 133 | 17 | Q8ZLJ3 | Q8ZLJ3 pyrobaculum |
| 30 | 37 | 59.7 | 181 | 12 | Q80BT9 | Q80BT9 spodoptera |
| 31 | 37 | 59.7 | 266 | 2 | Q9ANW9 | Q9ANW9 bacteroides |
| 32 | 37 | 59.7 | 306 | 16 | Q83MF7 | Q83MF7 shigella fl |
| 33 | 37 | 59.7 | 306 | 16 | Q7UDS4 | Q7UDS4 shigella fl |
| 34 | 37 | 59.7 | 330 | 16 | Q89EI3 | Q89EI3 bradyrhizob |
| 35 | 37 | 59.7 | 333 | 13 | Q9PTW5 | Q9PTW5 paralicthy |
| 36 | 37 | 59.7 | 333 | 13 | Q8AWX8 | Q8AWX8 gadus morhu |
| 37 | 37 | 59.7 | 334 | 13 | Q90Z48 | Q90Z48 oncornychu |
| 38 | 37 | 59.7 | 334 | 13 | Q90ZF1 | Q90ZF1 oncornychu |
| 39 | 37 | 59.7 | 372 | 5 | Q18043 | Q18043 caenorhabdi |
| 40 | 37 | 59.7 | 606 | 2 | Q9AM03 | Q9AM03 helicobacte |
| 41 | 37 | 59.7 | 606 | 16 | Q9ZK94 | Q9ZK94 helicobacte |
| 42 | 37 | 59.7 | 606 | 16 | Q25744 | Q25744 helicobacte |
| 43 | 37 | 59.7 | 624 | 9 | Q85821 | Q85821 bacterioph |
| 44 | 37 | 59.7 | 642 | 16 | Q9XZ68 | Q9XZ68 thermotoga |
| 45 | 37 | 59.7 | 680 | 16 | Q88552 | Q88552 pseudomonas |

ALIGNMENTS

RESULT 1
Q813Q3
ID Q813Q3 PRELIMINARY; PRT; 2054 AA.
AC Q813Q3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PFE1045C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=36329;
RX (1) SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden P., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
Humphrey S., Jagels K., James K.D., Johnson D., Kerhornou A.,
Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
Sulston J.E., Craig A., Newbold C., Barrell B.G;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531(2002).
DR EMBL; AL929353; CAD51573.1; -.
KW Hypothetical protein.

```

SQ SEQUENCE 2054 AA; 244096 MW; 522D7E233ACD1D1A CRC64;
Query Match 71.0%; Score 44; DB 5; Length 2054;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
Db 1046 LTYIDINVDY 1055

RESULT 2
Q8P041 PRELIMINARY; PRT; 259 AA.
ID Q8P041
AC Q8P041
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein spym18_1578.
GN SPYM18_1578.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C.; Barbican K.D.; Van Gompel J.J.; Smoot L.M.; Chaussee M.S.;
RA Sylva G.L.; Sturdevant D.E.; Ricklefs S.W.; Porcella S.F.;
RA Parkins L.D.; Beres S.B.; Campbell D.S.; Smith T.M.; Zhang Q.;
RA Kapur V.; Daly J.A.; Veasy L.G.; Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; A010071; AAU98141.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 259 AA; 30329 MW; 40790BA21E3AC18B CRC64;

Query Match 69.4%; Score 43; DB 16; Length 259;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
Db 150 ITRYDYEFVY 159

RESULT 3
Q8K6N0 PRELIMINARY; PRT; 259 AA.
ID Q8K6N0
AC Q8K6N0
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein spym3_1272.
GN SPYM3_1272 OR SPS0590.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B.; Sylva G.L.; Barbican K.D.; Lei B.; Hoff J.S.;
RA Mammarella N.D.; Liu M.-Y.; Smoot J.C.; Porcella S.F.; Parkins L.D.;
RA Campbell D.S.; Smith T.M.; McCormick J.K.; Leung D.Y.M.;
RA Schlievert P.M.; Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";

Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
[2]
SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I.; Kurokawa K.; Nakata M.; Tomiyasu Y.; Yamashita A.;
RA Yamazaki K.; Okahashi N.; Kawabata S.; Yasunaga T.; Hattori M.;
RA Hayashi H.; Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A014161; AM79879.1; -
DR EMBL; AP005142; BAC6385.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 259 AA; 30337 MW; 70B8BEE221C8D90 CRC64;

Query Match 69.4%; Score 43; DB 16; Length 259;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
Db 150 ITRYDYEFVY 159

RESULT 4
Q97WY9 PRELIMINARY; PRT; 130 AA.
ID Q97WY9
AC Q97WY9;
DT 01-OCT-2001 (T-EMBLrel. 18, Created)
DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein SSO1970.
GN SSO1970.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OC NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q.; Singh R.K.; Confalonieri F.; Zivanovic Y.; Allard G.;
RA Awavey M.J.; Chan-Weiher C.C.-Y.; Clausen I.G.; Curtis B.A.;
RA De Moers A.; Etrauso G.; Fletcher C.; Gordon P.M.K.;
RA Heikamp-de Jong I.; Jeffries A.C.; Kozera C.J.; Medina N.; Peng X.;
RA Thi-Ngoc H.P.; Redder P.; Schenk M.E.; Theriault C.; Tolstrup N.;
RA Charlebois R.L.; Doolittle W.F.; Duguet M.; Gaasterland T.;
RA Garrett R.A.; Ragan M.A.; Sensen C.W.; Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; A006804; AAK42162.1; -
DR PIR; C90363; C90363.
DR InterPro; IPR002716; PIN.
DR InterPro; IPR00596; PINC.
DR Pfam; PF01850; PIN; 1.
DR SMART; SMO0670; PINC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 130 AA; 15106 MW; 5B0B966B366BD159 CRC64;

Query Match 66.1%; Score 41; DB 17; Length 130;
Best Local Similarity 70.0%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
Db 89 LTYIDASYAY 98

RESULT 5
Q8E3E1 PRELIMINARY; PRT; 210 AA.
ID Q8E3E1
AC Q8E3E1;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)

```

```

Query Match          66.1%; Score 41; DB 16; Length 210;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 LTYDYERAY 11
DB      12 LTYFYDFDY 21
      |||::|||
      |||::|||

RESULT 7
Q860W8 PRELIMINARY; PRT; 323 AA.
AC Q860W8;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Major histocompatibility complex class Ib M10.1 (Fragment).
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J;
RX MEDLINE=22515520; PubMed=12628182;
RA Locanto J., Papes F., Chang E., Stowers L., Jones E.F., Takada T.,
RA Kamanovics A., Fischer Lindahl K., Dulac C.;
RA "Functional Expression of Murine V2R Pheromone Receptors Involves
RT Selective Association with the M10 and M1 Families of MHC Class Ib
RT Molecules.";
RL Cell 112:607-618(2003).
DR EMBL; AY211078; AAC42745.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MEC1.1.
DR PRINIS; PRO1638; MHCCLASSI.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
DR PROSITE; PS0290; IG_MHC; 1.
DR NON_TER 1
SQ SEQUENCE 323 AA; 37016 MW; CDB8GC1D816828 CRC64;

Query Match          64.5%; Score 40; DB 7; Length 323;
Best Local Similarity 87.5%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 LTYDYEF 9
DB      132 LTYGYEF 139
      |||||
      |||||

RESULT 8
Q85ZW9 PRELIMINARY; PRT; 329 AA.
AC Q85ZW9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Histocompatibility 2, M region locus 10.2.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RX MEDLINE=22507899; PubMed=12620187;

```

```

RA Ishii T., Hirota J., Mombaerts P.;
RT "Combinatorial Coexpression of Neural and Immune Multigene Families in
RL Mouse Vomeronasal Sensory Neurons.";
RL Curr. Biol. 13:394-400(2003).
DR EMBL; AF539598; AAC050321.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I.
DR PRINTS; PR01638; MHCCLASSI.
DR SMART; SM00407; ICG1; 1.
DR PROSITE; PS00835; Ig_LIKE; 1.
DR PROSITE; PS00290; Ig_MHC; 1.
SQ SEQUENCE 329 AA; 37517 MW; B70DB5CAA34E422E CRC64;

Query Match 64.5%; Score 40; DB 7; Length 329;
Best Local Similarity 87.5%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTYDYEF 9
Db 138 LTYGYEF 145

RESULT 9
Q857E5 PRELIMINARY; PRT; 77 AA.
AC Q857E5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Q877.
OS Mycobacteriophage Bx22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205870;
RN [1]
RP MEDLINE=22592650; PubMed=12705866;
RA Pedulla M.L., Ford M.B., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandaamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AV129332; AN01831.1; -.
SQ SEQUENCE 77 AA; 8809 MW; B8D295DDB02C82ED6 CRC64;

Query Match 62.9%; Score 39; DB 9; Length 77;
Best Local Similarity 63.6%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
Db 48 AITDYNELAY 58

RESULT 10
Q9CV82 PRELIMINARY; PRT; 495 AA.
AC Q9CV82;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 2310003L06RIK protein (Fragment).
GN 2310003L06RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/60; Tissue=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009122; BAB26087.1; -.
DR MGD; MGI:1921498; 2310003L06RIK.
FT NON TER 495 495
SQ SEQUENCE 495 AA; 55471 MW; 1FE76D923CEDE9E7 CRC64;

Query Match 62.9%; Score 39; DB 11; Length 495;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALTYDYEF 9
Db 171 SVTYQYEF 179

RESULT 11
Q89YL8 PRELIMINARY; PRT; 580 AA.
AC Q89YL8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN BT4713.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550859; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016946; AAQ79818.1; -.
DR GO; GO:0003743; P:translation initiation factor activity; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR008979; Gal_bind_like.
DR PRODOM; PD004078; eIF5_eIF2B; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 580 AA; 65978 MW; E6B0C8B3017922E CRC64;

Query Match 62.9%; Score 39; DB 16; Length 580;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 3 TYDYEFAY 11
| | | | |
Db 162 TYDYPRAY 170

RESULT 12

Q9EMU7 Q9EMU7 PRELIMINARY; PRT; 148 AA.
AC Q9EMU7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE AMV102.
GN AMV102.
OS Amsacta moorei entomopoxvirus (AmEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxID=28321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20396580; PubMed=10936094;
RA Bowden A.L., Glassberg K.J., Diggins J., Shaw R., Farmerie W.,
RA Moyer R.W.;
RT "Complete Genomic Sequence of the Amsacta moorei Entomopoxvirus:
RT Analysis and Comparison with Other Poxviruses.";
RL Virology 274:120-139(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Bowden A.L., Glassberg K.J., Diggins J., Shaw R., Farmerie W.,
RA Moyer R.W.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF250284; AAC02808.1; -
SQ SEQUENCE 148 AA; 18636 MW; DAA4C41A79AF04D2 CRC64;

Query Match 61.3%; Score 38; DB 12; Length 148;
Best Local Similarity 50.0%; Pred. No. 86;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
: | | | | |
Db 137 MEYDYKYIY 146

RESULT 13

Q8XLE4 Q8XLE4 PRELIMINARY; PRT; 154 AA.
AC Q8XLE4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CPE1098.
GN CPE1098.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
RL EMBL: AP003189; BAB80804.1; -
KN Hypothetical protein; Complete proteome.
SQ SEQUENCE 154 AA; 17870 MW; B5370264C4F9B7A6 CRC64;

Query Match 61.3%; Score 38; DB 16; Length 154;
Best Local Similarity 70.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
| | | | |
Db 46 LTYDYELDY 55

RESULT 14

Q8P7X5 Q8P7X5 PRELIMINARY; PRT; 255 AA.
AC Q8P7X5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE VirB9 protein.
GN VIRB9 OR XCC2480.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;

"Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RT Nature 417:459-463(2002).
DR EMBL: AE012359; AAM41756.1; -
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004357; IVSec_cagX.
DR Pfam; PF03524; cagX; 1.
KW Complete proteome.
SQ SEQUENCE 255 AA; 29205 MW; FDC8C0570E8871C CRC64;

Query Match 61.3%; Score 38; DB 16; Length 255;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 10
| | | | |
Db 162 LTYDYDYA 170

RESULT 15

Q8JGW5 Q8JGW5 PRELIMINARY; PRT; 333 AA.
AC Q8JGW5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)
DE (GAPDH).
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Peinado J.R., Castano J.P., Sanchez-Hormigo A., Anouar Y., Tonon M.C.,

RA Vaudry H., Gracia-Navarro F., Malagon M.M.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
 CC -I- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + PHOSPHATE +
 CC NAD(+) = 3-PHOSPHO-D-GLYCEROYL PHOSPHATE + NADH.
 CC -I- PATHWAY: SECOND PHASE OF GLYCOLYSIS: FIRST STEP.
 CC -I- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR ENBL; AY072703; AAL62488.1; -.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Glycolysis; NAD; Oxidoreductase.
 SQ SEQUENCE 333 AA; 35679 MW; 0A93AA634C6974FC CRC64;

Query Match 61.3%; Score 38; DB 13; Length 333;
 Best Local Similarity 60.0%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
 DB 309 ITWYDNEFGY 318

Search completed: October 6, 2004, 16:33:52
 Job time : 41.8596 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 68.5088 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-6

Perfect score: 62

Sequence: 1 ALTYYDYEFAY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 62 | 100.0 | 11 | 10 | US-09-798-689-30 |
| 2 | 62 | 100.0 | 11 | 10 | US-09-996-954B-6 |
| 3 | 62 | 100.0 | 11 | 12 | US-10-374-600-111 |
| 4 | 62 | 100.0 | 11 | 15 | US-10-374-531-111 |
| 5 | 62 | 100.0 | 13 | 12 | US-10-374-600-99 |
| 6 | 62 | 100.0 | 13 | 15 | US-10-374-531-99 |
| 7 | 62 | 100.0 | 119 | 12 | US-10-374-600-115 |
| 8 | 62 | 100.0 | 119 | 12 | US-10-374-600-116 |
| 9 | 62 | 100.0 | 119 | 12 | US-10-374-600-117 |
| 10 | 62 | 100.0 | 119 | 12 | US-10-374-600-118 |
| 11 | 62 | 100.0 | 119 | 12 | US-10-374-600-119 |
| 12 | 62 | 100.0 | 119 | 12 | US-10-374-600-120 |
| 13 | 62 | 100.0 | 119 | 15 | US-10-374-531-115 |
| 14 | 62 | 100.0 | 119 | 15 | US-10-374-531-116 |
| 15 | 62 | 100.0 | 119 | 15 | US-10-374-531-117 |

16 62 100.0 119 15 US-10-374-531-118 Sequence 118, App
17 62 100.0 119 15 US-10-374-531-119 Sequence 119, App
18 62 100.0 119 15 US-10-374-531-120 Sequence 120, App
19 62 100.0 138 12 US-10-374-600-8 Sequence 8, Appli
20 62 100.0 138 12 US-10-374-600-13 Sequence 13, Appli
21 62 100.0 138 12 US-10-374-600-19 Sequence 19, Appli
22 62 100.0 138 15 US-10-374-531-8 Sequence 8, Appli
23 62 100.0 138 15 US-10-374-531-13 Sequence 13, Appli
24 62 100.0 138 15 US-10-374-531-19 Sequence 19, Appli
25 48 77.4 132 12 US-10-374-600-24 Sequence 24, Appli
26 48 77.4 132 12 US-10-374-600-26 Sequence 26, Appli
27 48 77.4 132 12 US-10-374-600-27 Sequence 27, Appli
28 48 77.4 132 12 US-10-374-600-28 Sequence 28, Appli
29 48 77.4 132 12 US-10-374-600-29 Sequence 29, Appli
30 48 77.4 132 12 US-10-374-600-30 Sequence 30, Appli
31 48 77.4 132 15 US-10-374-531-24 Sequence 24, Appli
32 48 77.4 132 15 US-10-374-531-26 Sequence 26, Appli
33 48 77.4 132 15 US-10-374-531-27 Sequence 27, Appli
34 48 77.4 132 15 US-10-374-531-28 Sequence 28, Appli
35 48 77.4 132 15 US-10-374-531-29 Sequence 29, Appli
36 48 77.4 132 15 US-10-374-531-30 Sequence 30, Appli
37 40 64.5 118 9 US-09-905-243-78 Sequence 78, Appli
38 40 64.5 119 9 US-09-905-243-79 Sequence 79, Appli
39 38 61.3 292 16 US-10-767-701-41987 Sequence 41987, A
40 38 61.3 599 16 US-10-437-963-116912 Sequence 116912, A
41 38 61.3 815 15 US-10-104-047-2496 Sequence 2496, Ap
42 37 59.7 13 9 US-09-839-666-2 Sequence 2, Appli
43 37 59.7 13 12 US-10-372-735-44 Sequence 44, Appli
44 37 59.7 13 14 US-10-234-579-2 Sequence 2, Appli
45 37 59.7 14 9 US-09-839-666-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-798-689-30
; Sequence 30, Application US/09798689
; Publication No. US20030103973A1
; GENERAL INFORMATION:
; APPLICANT: Rockwell, Patricia
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: Method for Reducing Tumor Growth with VEGF Antagonists
; TITLE OF INVENTION: Combined With Radiation and Chemotherapy
; FILE REFERENCE: Sequence Listings 1-41 for 381-25 CIP
; CURRENT APPLICATION NUMBER: US/09/798,689
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 03/401,163
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/967,113
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 08/706,804
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/476,533
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/326,552
; PRIOR FILING DATE: 1994-10-20
; PRIOR APPLICATION NUMBER: 08/196,041
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Mouse
US-09-798-689-30

Query Match 100.0%; Score 62; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYYDYEFAY 11
|||||||

Db 1 ALTYDYEFAY 11

RESULT 2

US-09-996-954B-6

; Sequence 6, Application US/09996954B

; Publication No. US20030157104A1

; GENERAL INFORMATION:

; APPLICANT: Mksal, Harlan W.

; TITLE OF INVENTION: Treatment of Refractory Human Tumors

; FILE OF INVENTION: With Epidermal Growth Factor Receptor Antagonists

; FILE REFERENCE: 11245-46605

; CURRENT APPLICATION NUMBER: US/09/996,954B

; CURRENT FILING DATE: 2001-11-30

; PRIOR APPLICATION NUMBER: 09/840,146

; PRIOR FILING DATE: 04-24-2001

; PRIOR APPLICATION NUMBER: 09/374,028

; PRIOR FILING DATE: 08-13-1999

; PRIOR APPLICATION NUMBER: 09/312,284

; PRIOR FILING DATE: 05-14-1999

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 11

; TYPE: PRT

; ORGANISM: Homo sapiens-Rodent Chimera

US-09-996-954B-6

Query Match 100.0%; Score 62; DB 10; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11

|||||

Db 1 ALTYDYEFAY 11

RESULT 3

US-10-374-600-111

; Sequence 111, Application US/10374600

; Publication No. US20030224001A1

; GENERAL INFORMATION:

; APPLICANT: Imclone Systems Incorporated, et al.

; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,600

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

US-09-996-954B-6

Query Match 100.0%; Score 62; DB 12; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11

|||||

Db 1 ALTYDYEFAY 11

RESULT 4

US-10-374-531-111

; Sequence 111, Application US/10374531

; Publication No. US20040006212A1

; GENERAL INFORMATION:

; APPLICANT: Imclone Systems Incorporated, et al.

; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,531

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

US-10-374-600-111

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 111:
US-10-374-531-111

Query Match 100.0%; Score 62; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYDEFAY 11

DB 1 ALTYDYDEFAY 11

RESULT 5

US-10-374-600-99
; Sequence 99, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:

APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,600

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 99:

US-10-374-600-99

Query Match 100.0%; Score 62; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYDEFAY 11

DB 2 ALTYDYDEFAY 12

RESULT 6

US-10-374-531-99
; Sequence 99, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:

APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: One Broadway

CITY: New York

STATE: New York

COUNTRY: US

ZIP: 10004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,531

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C

FILING DATE: 19-Mar-1998

APPLICATION NUMBER: PCT/US96/09847

FILING DATE: 07-JUN-1996

APPLICATION NUMBER: US 08/482,982

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/573,289

FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deborah A. Somerville

REGISTRATION NUMBER: 31,995

REFERENCE/DOCKET NUMBER: 11245/46003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 425-7200

TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 99:

US-10-374-531-99

Query Match 100.0%; Score 62; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYDEFAY 11

DB 2 ALTYDYDEFAY 12

RESULT 7

US-10-374-600-115
; Sequence 115, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:

APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR


```

; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-5288
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-10-374-600-117

Query Match 100.0%; Score 62; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYDFAY 11
Db 98 ALTYDYDFAY 108

RESULT 10
US-10-374-600-118
; Sequence 118, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-5288
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 119:

```

```

; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 118:
US-10-374-600-118

Query Match 100.0%; Score 62; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYDFAY 11
Db 98 ALTYDYDFAY 108

RESULT 11
US-10-374-600-119
; Sequence 119, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-5288
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 119:

```

```
US-10-374-600-119
Query Match      100.0%; Score 62; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALTYDYEFAY 11
      |||||
Db      98 ALTYDYEFAY 108

RESULT 12
US-10-374-600-120
; Sequence 120, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 120:
US-10-374-600-120
Query Match      100.0%; Score 62; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALTYDYEFAY 11
      |||||
Db      98 ALTYDYEFAY 108

US-10-374-531-115
Query Match      100.0%; Score 62; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALTYDYEFAY 11
      |||||
Db      98 ALTYDYEFAY 108

RESULT 13
US-10-374-531-115
; Sequence 115, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 115:
US-10-374-531-115
Query Match      100.0%; Score 62; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALTYDYEFAY 11
      |||||
Db      98 ALTYDYEFAY 108

RESULT 14
US-10-374-531-116
; Sequence 116, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
```

STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-7200
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-10-374-531-116.
Query Match 100.0%; Score 62; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 98 ALTYDYEFAY 108
RESULT 15
US-10-374-531-117
Sequence 117, Application US/10374531
Publication No. US20040006212A1
GENERAL INFORMATION:
APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-7200
TELEFAX: (212) 425-5288
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-10-374-531-117
Query Match 100.0%; Score 62; DB 15; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 98 ALTYDYEFAY 108
Search completed: October 6, 2004, 17:09:03
Job time : 68.5088 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 19.4912 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-6
Perfect score: 62
Sequence: 1 ALTYDYEFAY 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A COMB pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB pep.*
3: /cgn2_6/prodata/2/iaa/5A COMB pep.*
4: /cgn2_6/prodata/2/iaa/5B COMB pep.*
5: /cgn2_6/prodata/2/iaa/5A COMB pep.*
6: /cgn2_6/prodata/2/iaa/5B COMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 39 | 62.9 | 119 | 2 | US-08-800-198-2 |
| 2 | 39 | 62.9 | 119 | 3 | US-09-296-595-2 |
| 3 | 39 | 62.9 | 240 | 2 | US-08-800-198-8 |
| 4 | 39 | 62.9 | 240 | 3 | US-09-296-595-8 |
| 5 | 38 | 61.3 | 509 | 3 | US-08-369-8220-26 |
| 6 | 38 | 61.3 | 509 | 3 | US-08-582-7760-41 |
| 7 | 38 | 61.3 | 509 | 3 | US-08-434-8315-38 |
| 8 | 37 | 59.7 | 13 | 2 | US-08-737-085A-2 |
| 9 | 37 | 59.7 | 13 | 3 | US-09-246-258-2 |
| 10 | 37 | 59.7 | 13 | 3 | US-09-532-106-2 |
| 11 | 37 | 59.7 | 13 | 4 | US-09-839-666-2 |
| 12 | 37 | 59.7 | 14 | 2 | US-08-737-085A-1 |
| 13 | 37 | 59.7 | 14 | 3 | US-09-246-258-1 |
| 14 | 37 | 59.7 | 14 | 3 | US-09-532-106-1 |
| 15 | 37 | 59.7 | 14 | 4 | US-09-839-666-1 |
| 16 | 37 | 59.7 | 22 | 2 | US-08-737-085A-13 |
| 17 | 37 | 59.7 | 22 | 3 | US-09-246-258-13 |
| 18 | 37 | 59.7 | 22 | 3 | US-09-532-106-13 |
| 19 | 37 | 59.7 | 22 | 4 | US-09-839-666-13 |
| 20 | 37 | 59.7 | 24 | 2 | US-08-737-085A-14 |
| 21 | 37 | 59.7 | 24 | 2 | US-08-737-085A-17 |
| 22 | 37 | 59.7 | 24 | 2 | US-08-737-085A-18 |
| 23 | 37 | 59.7 | 24 | 2 | US-08-737-085A-23 |
| 24 | 37 | 59.7 | 24 | 3 | US-09-246-258-14 |
| 25 | 37 | 59.7 | 24 | 3 | US-09-246-258-17 |
| 26 | 37 | 59.7 | 24 | 3 | US-09-246-258-18 |
| 27 | 37 | 59.7 | 24 | 3 | US-09-246-258-23 |

| | | | | | | |
|----|----|------|----|---|-------------------|-------------------|
| 28 | 37 | 59.7 | 24 | 3 | US-09-532-106-14 | Sequence 14, Appl |
| 29 | 37 | 59.7 | 24 | 3 | US-09-532-106-17 | Sequence 17, Appl |
| 30 | 37 | 59.7 | 24 | 3 | US-09-532-106-18 | Sequence 18, Appl |
| 31 | 37 | 59.7 | 24 | 3 | US-09-532-106-23 | Sequence 23, Appl |
| 32 | 37 | 59.7 | 24 | 4 | US-09-839-666-14 | Sequence 14, Appl |
| 33 | 37 | 59.7 | 24 | 4 | US-09-839-666-17 | Sequence 17, Appl |
| 34 | 37 | 59.7 | 24 | 4 | US-09-839-666-18 | Sequence 18, Appl |
| 35 | 37 | 59.7 | 24 | 4 | US-09-839-666-23 | Sequence 23, Appl |
| 36 | 37 | 59.7 | 25 | 2 | US-08-737-085A-19 | Sequence 19, Appl |
| 37 | 37 | 59.7 | 25 | 3 | US-09-246-258-19 | Sequence 19, Appl |
| 38 | 37 | 59.7 | 25 | 3 | US-09-532-106-19 | Sequence 19, Appl |
| 39 | 37 | 59.7 | 25 | 4 | US-09-839-666-19 | Sequence 19, Appl |
| 40 | 37 | 59.7 | 26 | 2 | US-08-737-085A-15 | Sequence 15, Appl |
| 41 | 37 | 59.7 | 26 | 2 | US-08-737-085A-16 | Sequence 16, Appl |
| 42 | 37 | 59.7 | 26 | 3 | US-09-246-258-15 | Sequence 15, Appl |
| 43 | 37 | 59.7 | 26 | 3 | US-09-246-258-16 | Sequence 16, Appl |
| 44 | 37 | 59.7 | 26 | 3 | US-09-532-106-15 | Sequence 15, Appl |
| 45 | 37 | 59.7 | 26 | 3 | US-09-532-106-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1
US-08-800-198-2
; Sequence 2, Application US/08800198
; Patent No. 5942602
; GENERAL INFORMATION:
; APPLICANT: WELLS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VAKALOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,198
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HAMLET-KING, DIANA
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SCH 1576
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-800-198-2

Query Match 62.9% ; Score 39; DB 2; Length 119;
Best Local Similarity 75.0% ; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY      4 YYDYEFAY 11
Db      101 YGYDFAY 108

RESULT 2
US-09-296-595-2
; Sequence 2, Application US/09296595A
; Patent No. 6129915
; GENERAL INFORMATION:
; APPLICANT: WELLS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VAKALOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; FILE REFERENCE: SCH-1576 D1
; CURRENT APPLICATION NUMBER: US/09/296,595A
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 08/800,198
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-296-595-2

Query Match      62.9%; Score 39; DB 3; Length 119;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 YYDYEFAY 11
Db      101 YGYDFAY 108

RESULT 3
US-08-800-198-8
; Sequence 8, Application US/08800198
; Patent No. 5942602
; GENERAL INFORMATION:
; APPLICANT: WELLS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VAKALOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,198
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HAMLET-KING, DIANA
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SCH 1576
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
US-08-800-198-8

Query Match      62.9%; Score 39; DB 2; Length 240;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 YYDYEFAY 11
Db      101 YGYDFAY 108

RESULT 4
US-09-296-595-8
; Sequence 8, Application US/09296595A
; Patent No. 6129915
; GENERAL INFORMATION:
; APPLICANT: WELLS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VAKALOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; FILE REFERENCE: SCH-1576 D1
; CURRENT APPLICATION NUMBER: US/09/296,595A
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 08/800,198
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-296-595-8

Query Match      62.9%; Score 39; DB 3; Length 240;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 YYDYEFAY 11
Db      101 YGYDFAY 108

RESULT 5
US-08-369-822C-26
; Sequence 26, Application US/08369822C
; Patent No. 6015660
; GENERAL INFORMATION:
; APPLICANT: Lipkin, W. I.
; APPLICANT: Briese, Thomas
; APPLICANT: Kliche, Stefanie
; APPLICANT: Schneider, Patrick A.
; APPLICANT: Stitz, Lothar
; APPLICANT: Schneemann, Anette
; TITLE OF INVENTION: Borna Disease Viral Sequences,
; TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
; TITLE OF INVENTION: System Diseases
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 865 South Figueroa Street, 29th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017-2571
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS NT-WORDPERFECT 8.0
SOFTWARE: ASCII (DOS) TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,822C
FILING DATE: 06-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Churchill, Margaret A. (Ph.D.)
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1279-194XX
TELEPHONE: 213/892-9200
TELEFAX: 213/680-4518
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-369-822C-26

Query Match 61.3%; Score 38; DB 3; Length 509;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
||:||||:
Db 446 TYDSEFSY 454

RESULT 6
US-08-582-776C-41
Sequence 41, Application US/08582776C
Patent No. 6077510
GENERAL INFORMATION:
APPLICANT: Lipkin, W. I.
APPLICANT: Briese, Thomas
APPLICANT: Kliche, Stefanie
APPLICANT: Schneider, Patrick A.
APPLICANT: Stitz, Lothar
APPLICANT: Schneemann, Anette
TITLE OF INVENTION: Borna Disease Viral Sequences,
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
TITLE OF INVENTION: System Diseases
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 865 South Figueroa Street, 29th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017-2576
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS NT
SOFTWARE: ASCII DOS TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/582,776C
FILING DATE: 04-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/369,822
FILING DATE: 06-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,831
FILING DATE: 04-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Churchill, Margaret A.

REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1279-194C2
TELEPHONE: 213/892-9200
TELEFAX: 213/680-4518
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-582-776C-41

Query Match 61.3%; Score 38; DB 3; Length 509;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 TYDYEFAY 11
||:||||:
Db 446 TYDSEFSY 454

RESULT 7
US-08-434-831B-38
Sequence 38, Application US/08434831B
Patent No. 6113905
GENERAL INFORMATION:
APPLICANT: Lipkin, W. I.
APPLICANT: Briese, Thomas
APPLICANT: Kliche, Stefanie
APPLICANT: Schneider, Patrick A.
APPLICANT: Stitz, Lothar
APPLICANT: Schneemann, Anette
TITLE OF INVENTION: Borna Disease Viral Sequences,
TITLE OF INVENTION: Diagnostics and Therapeutics for Central Nervous
TITLE OF INVENTION: System Diseases
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 865 South Figueroa Street, 29th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017-2571
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS NT-WORDPERFECT 8.0
SOFTWARE: ASCII (DOS) TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,831B
FILING DATE: 04-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/369,822
FILING DATE: 06-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Churchill, Margaret A. (Ph.D.)
REGISTRATION NUMBER: 39,944
REFERENCE/DOCKET NUMBER: 1279-194C1
TELEPHONE: 213/892-9200
TELEFAX: 213/680-4518
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-434-831B-38

Query Match 61.3%; Score 38; DB 3; Length 509;
Best Local Similarity 66.7%; Pred. No. 81;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LTYDYEFAY 11
Db 446 TYFDSEFSY 454

RESULT 8

US-08-737-085A-2
; Sequence 2, Application US/08737085A
; Patent No. 5869232
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,085A
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 426
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/OC569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-737-085A-2

Query Match 59.7%; Score 37; DB 2; Length 13;
Best Local Similarity 70.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LTYDYEFAY 11
Db 3 LIYDYEDDY 12

US-09-246-258-2
; Sequence 2, Application US/09246258
; Patent No. 6040137
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER

RESULT 9

US-09-246-258-2
; Sequence 2, Application US/09246258
; Patent No. 6040137
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER

US-09-532-106-2
; Sequence 2, Application US/09532106
; Patent No. 6245895
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/246,258
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/737,085
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/OC569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-246-258-2

Query Match 59.7%; Score 37; DB 3; Length 13;
Best Local Similarity 70.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LTYDYEFAY 11
Db 3 LIYDYEDDY 12

RESULT 10

US-09-532-106-2
; Sequence 2, Application US/09532106
; Patent No. 6245895
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/532,106
; FILING DATE: 21-Mar-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,085A
; FILING DATE: 27-DEC-1996

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/0C569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-532-106-2

Query Match 59.7%; Score 37; DB 3; Length 13;
Best Local Similarity 70.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
DB 3 LIYDYEDY 12

RESULT 11
US-09-839-666-2
; Sequence 2, Application US/09839666
; Patent No. 6469143
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/737,085
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/0C569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-839-666-2

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/0C569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-737-085A-1

Query Match 59.7%; Score 37; DB 2; Length 14;
Best Local Similarity 70.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
DB 3 LIYDYEDY 12

RESULT 12
US-08-737-085A-1
; Sequence 1, Application US/08737085A
; Patent No. 5869232
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,085A
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 426
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/0C569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-737-085A-1

Query Match 59.7%; Score 37; DB 2; Length 14;
Best Local Similarity 70.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11
DB 3 LIYDYEDY 12

RESULT 13
US-09-246-258-1
; Sequence 1, Application US/09246258
; Patent No. 6040137
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

```


Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LTYDYEFAY 11

Db 3 LTYDYEDY 12

Search completed: October 6, 2004, 16:36:32
Job time : 20.4912 secs

This Page Blank (uspio)

02-3

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 55.5789 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-6

Sequence: 1 ALTYDYEFAY 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A Genesep29Jan04:*
1: Genesep21980s:*
2: Genesep21990s:*
3: Genesep22000s:*
4: Genesep22001s:*
5: Genesep22002s:*
6: Genesep22003as:*
7: Genesep22003bs:*
8: Genesep22004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|-------------|
| 1 | 62 | 100.0 | 11 | 3 AAY59313 | Heavy cha |
| 2 | 62 | 100.0 | 11 | 4 AAB37954 | Anti-EGFR |
| 3 | 62 | 100.0 | 11 | 5 AAU77786 | Mouse hea |
| 4 | 62 | 100.0 | 119 | 2 AAW08953 | Heavy cha |
| 5 | 62 | 100.0 | 119 | 2 AAW08950 | Heavy cha |
| 6 | 62 | 100.0 | 119 | 2 AAW08952 | Heavy cha |
| 7 | 62 | 100.0 | 119 | 2 AAW08951 | Heavy cha |
| 8 | 62 | 100.0 | 119 | 2 AAW08954 | Heavy cha |
| 9 | 62 | 100.0 | 138 | 2 AAW08947 | Heavy cha |
| 10 | 62 | 100.0 | 138 | 2 AAW08944 | Heavy cha |
| 11 | 62 | 100.0 | 138 | 2 AAW08942 | Heavy cha |
| 12 | 62 | 100.0 | 240 | 2 AAW05133 | Single ch |
| 13 | 62 | 100.0 | 651 | 2 AAW05135 | scFv (225) |
| 14 | 62 | 100.0 | 892 | 2 AAW05140 | scFv (225) |
| 15 | 62 | 100.0 | 892 | 2 AAW05139 | scFv2 (FRP |
| 16 | 62 | 100.0 | 1020 | 2 AAW05141 | scFv2 (FRP |
| 17 | 42 | 67.7 | 993 | 4 ABG10386 | Novel hum |
| 18 | 41 | 66.1 | 210 | 5 ABG28261 | Streptoco |
| 19 | 40 | 64.5 | 10 | 3 AAB19756 | Erythropro |
| 20 | 40 | 64.5 | 118 | 3 AAY56681 | Anti-eryt |
| 21 | 40 | 64.5 | 119 | 3 AAY56682 | Humanized |
| 22 | 40 | 64.5 | 120 | 3 AAY56705 | Murine an |
| 23 | 40 | 64.5 | 120 | 3 AAB19745 | Erythropro |
| 24 | 40 | 64.5 | 125 | 3 AAB19748 | Erythropro |
| 25 | 40 | 64.5 | 125 | 3 AAB19747 | Erythropro |

| | | | | | |
|----|------|------|-----|------------|-----------|
| 26 | 39 | 62.9 | 119 | 2 AAW71240 | Heavy cha |
| 27 | 39 | 62.9 | 240 | 2 AAW71243 | scFv comp |
| 28 | 38.5 | 62.1 | 120 | 2 AAR13721 | Control f |
| 29 | 38 | 61.3 | 322 | 7 ABM74082 | DNA clone |
| 30 | 38 | 61.3 | 471 | 7 ABM73723 | DNA clone |
| 31 | 38 | 61.3 | 561 | 6 ABP75850 | Human sec |
| 32 | 38 | 61.3 | 815 | 7 ADB64342 | Human pro |
| 33 | 37 | 59.7 | 13 | 2 AAR84527 | CDR3 of |
| 34 | 37 | 59.7 | 14 | 2 AAR84526 | CDR3 of |
| 35 | 37 | 59.7 | 14 | 2 AAW38244 | CDR3-H3 s |
| 36 | 37 | 59.7 | 22 | 2 AAR84538 | Anti-HIV- |
| 37 | 37 | 59.7 | 24 | 2 AAR84539 | Anti-HIV- |
| 38 | 37 | 59.7 | 24 | 2 AAR84543 | Anti-HIV- |
| 39 | 37 | 59.7 | 24 | 2 AAR84548 | Anti-HIV- |
| 40 | 37 | 59.7 | 24 | 2 AAR84542 | Anti-HIV- |
| 41 | 37 | 59.7 | 25 | 2 AAR84544 | 2[Anti-HI |
| 42 | 37 | 59.7 | 26 | 2 AAR84540 | Anti-HIV- |
| 43 | 37 | 59.7 | 26 | 2 AAR84541 | Anti-HIV- |
| 44 | 37 | 59.7 | 115 | 4 ABG03953 | Novel hum |
| 45 | 37 | 59.7 | 120 | 2 AAR48619 | Sequence |

ALIGNMENTS

RESULT 1
AAY59313
ID AAY59313 standard; peptide; 11 AA.

XX AC AAY59313;
XX DT 07-MAR-2000 (first entry)
XX DE Heavy chain hypervariable region, CDR3.
XX KW Hypervariable region; complementarity determining region; CDR; tumour;
XX KW single chain antibody; growth inhibitor; human; tumourigenesis; therapy;
XX KW protein receptor tyrosine kinase; heavy chain.

OS Mus sp.
XX PN WO9960023-A1.
XX PD 25-NOV-1999.
XX PF 14-MAY-1999; 99WO-US010741.
XX PR 15-MAY-1998; 98US-00079612.
XX PR 15-MAY-1998; 98US-0085613P.
XX PR 07-DEC-1998; 98US-00206138.
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (UABR-) UAB RES FOUND.
XX PI Waksal HW, Saleh MN, Robert F, Buchsbaum DJ;
XX DR WPI; 2000-062440/05.
XX DR N-FSDB; AAZ48627.
XX PT Treatment of human tumors, using a combination of radiation and a non-
XX PT radiolabelled protein receptor tyrosine kinase inhibitor.
XX PS Disclosure; Page 15; 31pp; English.

CC This sequence is the hypervariable region CDR3 (complementarity
CC determining region 3) of the heavy chain of a single chain antibody
CC derived from the murine antibody 225. The invention relates to a method
CC for inhibiting the growth of tumours in human patients by treating with
CC an effective amount of a combination of radiation and a non-radiolabelled
CC protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
CC which can lead to tumourigenesis. The method can be used in the treatment
CC of tumours of e.g. breast, lung, colon, kidney, bladder, head and neck,
CC ovary, prostate or brain. The administration of a suitable antibody to

CC the patient makes the tumour more susceptible to radiotherapy

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
| | | | | | | | | |
DB 1 ALTYDYEFAY 11

RESULT 2
AAB37954
ID AAB37954 standard; protein; 11 AA.

XX AC AAB37954;
XX DT 12-MAR-2001 (first entry)
XX DE Anti-EGFR monoclonal antibody H chain V region CDR3 peptide sequence.

XX KW Refractory tumour growth inhibition; epidermal growth factor receptor;
KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
XX complementarity determining region; CDR.
XX OS Mus sp.
XX PN WO200069459-A1.
XX PD 23-NOV-2000.
XX PF 01-MAY-2000; 2000WO-US011756.
XX PR 14-MAY-1999; 99US-00312284.
XX PR 13-AUG-1999; 99US-00374028.
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PI Wakeal HW;
XX WPI; 2001-016160/02.
XX N-PSDB; AAC83237.

XX PT Epidermal growth factor receptor/human epidermal growth factor receptor-1
XX antagonist for inhibiting the growth of refractory tumors.

XX PS Disclosure; Page 13; 31pp; English.

XX CC This invention relates to a method for inhibiting the growth of
XX refractory tumors that are stimulated by a ligand of epidermal growth
XX factor receptor (EGFR) in human patients. The method involves treating
XX the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
XX optionally with a chemotherapeutic agent or radiation. The antagonist can
XX be for example a chimeric anti-EGFR monoclonal antibody, C225. The
XX EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
XX tumors such as tumors of breast, heart, lung, small intestine, colon,
XX spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas,
XX skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
XX liver, preferably squamous cell carcinomas. The present sequence
XX represents the heavy chain variable region complementarity determining
XX region 3 amino acid sequence of the chimeric anti-EGFR monoclonal
XX antibody C225 which is used in an example illustrating the method of the
XX invention

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11

DB 1 ALTYDYEFAY 11
| | | | | | | | | |

RESULT 3
AAU77786
ID AAU77786 standard; peptide; 11 AA.

XX AC AAU77786;
XX DT 05-JUN-2002 (first entry)
XX DE Mouse heavy chain hypervariable region (CDR3) of 225 antibody.

XX KW Mouse; heavy chain; antibody; hyperproliferative disease;
KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
KW seborrheic keratosis; warts; keloid scars; eczema; 255 antibody;
XX hypervariable region; CDR3; EGFR inhibitor.

XX OS Mus sp.
XX PN WO200211677-A2.
XX PD 14-FEB-2002.
XX PF 09-AUG-2001; 2001WO-US041647.
XX PR 09-AUG-2000; 2000US-00635974.
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PI Teufel T;
XX WPI; 2002-257423/30.
XX N-PSDB; ABK11442.

XX PT Treating a mammal with hyperproliferative disease especially psoriasis,
XX stimulated by ligand of member of epidermal growth factor family of
XX receptors, by administering antagonist of the receptor.

XX PS Disclosure; Page 12; 28pp; English.

XX CC This invention relates to a novel method for treating a mammal with
XX hyperproliferative disease stimulated by a ligand of a member of the
XX epidermal growth factor (EGF) family of receptors. The method involves
XX administering an antibody or a defective receptor that is an antagonist
XX of a member of the EGF receptor family, or a combination of the
XX antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
XX The antibody used in the method of the invention acts as an epidermal
XX growth factor receptor (EGFR) antagonist by inhibiting EGFR/HER1
XX phosphorylation. The method of the invention is useful for treating a
XX mammal with hyperproliferative disease such as psoriasis, actinic
XX keratosis, seborrheic keratosis, warts, keloid scars and eczema
XX stimulated by a ligand of a member of the EGF family of receptor. This
XX sequence represents the murine anti-EGF 255 antibody heavy chain (VH)
XX hypervariable region (CDR3) used as an inhibitor of EGFR in the method of
XX the invention

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 62; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
| | | | | | | | | |
DB 1 ALTYDYEFAY 11

RESULT 4
AAW08953
ID AAW08953 standard; protein; 119 AA.

XX

copying CDR-3

```

AC AAW08953;
XX
DT 18-SEP-1997 (first entry)
XX
DE Heavy chain variable region of 225RD antibody.
XX
KW Heavy chain; reshaped; monoclonal; antibody; 225RD; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
DR Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PS Claim 30; Fig 22; 112pp; English.
XX
CC The present sequence is the heavy chain variable region of the reshaped
CC human monoclonal antibody (MAB) H225, 225RHD. The MAB is specific for the
CC human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
CC be used to inhibit the growth of tumour cells, especially late stage
CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
CC agent, especially doxorubicin, taxol or cisplatin, or a signal
CC transduction, ras or cell cycle inhibitor
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 62; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
DB 98 ALTYDYEFAY 108

RESULT 5
AAW08950
ID AAW08950 standard; protein; 119 AA.
XX

```

```

AC AAW08950;
XX
DT 18-SEP-1997 (first entry)
XX
DE Heavy chain variable region of 225RA antibody.
XX
KW Heavy chain; reshaped; monoclonal; antibody; 225RA; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
DR Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PS Claim 30; Fig 22; 112pp; English.
XX
CC The present sequence is the heavy chain variable region of the reshaped
CC human monoclonal antibody (MAB) H225, 225RHA. The MAB is specific for the
CC human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
CC be used to inhibit the growth of tumour cells, especially late stage
CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
CC agent, especially doxorubicin, taxol or cisplatin, or a signal
CC transduction, ras or cell cycle inhibitor
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 62; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
DB 98 ALTYDYEFAY 108

RESULT 6
AAW08952
ID AAW08952 standard; protein; 119 AA.
XX

```

```

AC AAW08952;
XX
XX DT 18-SEP-1997 (first entry)
XX DE Heavy chain variable region of 225RC antibody.
XX
XX Heavy chain; reshaped; monoclonal; antibody; 225RC; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX
XX WO9640210-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 07-JUN-1996; 96WO-US009847.
XX
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Claim 30; Fig 22; 112pp; English.
XX
XX The present sequence is the heavy chain variable region of the reshaped
XX human monoclonal antibody (MAB) H225, 225RHC. The MAB is specific for the
XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX be used to inhibit the growth of tumour cells, especially late stage
XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX agent, especially doxorubicin, taxol or cisplatin, or a signal
XX transduction, ras or cell cycle inhibitor
XX
XX Sequence 119 AA;
XX
Query Match 100.0%; Score 62; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
DB 98 ALTYDYEFAY 108

RESULT 7
AAW08951
ID AAW08951 standard; protein; 119 AA.
XX

```

```

AC AAW08951;
XX
XX DT 18-SEP-1997 (first entry)
XX DE Heavy chain variable region of 225RB antibody.
XX
XX Heavy chain; reshaped; monoclonal; antibody; 225RB; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..30
FT /label= framework_1
FT Region 31..35
FT /label= CDR_1
FT Region 36..49
FT /label= framework_2
FT Region 50..65
FT /label= CDR_2
FT Region 66..97
FT /label= framework_3
FT Region 98..108
FT /label= CDR_3
FT Region 109..119
FT /label= framework_4
XX
XX WO9640210-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 07-JUN-1996; 96WO-US009847.
XX
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Claim 30; Fig 22; 112pp; English.
XX
XX The present sequence is the heavy chain variable region of the reshaped
XX human monoclonal antibody (MAB) H225, 225RHB. The MAB is specific for the
XX human epidermal growth factor (EGF) receptor. The MAB, or a fragment, can
XX be used to inhibit the growth of tumour cells, especially late stage
XX prostatic tumour cells in humans, optionally conjugated to a cytotoxic
XX agent, especially doxorubicin, taxol or cisplatin, or a signal
XX transduction, ras or cell cycle inhibitor
XX
XX Sequence 119 AA;
XX
Query Match 100.0%; Score 62; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALTYDYEFAY 11
DB 98 ALTYDYEFAY 108

RESULT 8
AAW08954
ID AAW08954 standard; protein; 119 AA.
XX

```

AAW08954;
18-SEP-1997 (first entry)
Heavy chain variable region of 225RE antibody.
Heavy chain; reshaped; monoclonal; antibody; 225RE; human;
epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
late stage; prostatic; prostate; variable region; framework;
complementarity determining region; CDR.
Homo sapiens.
Key Location/Qualifiers
Region 1..30
/label= framework_1
Region 31..35
/label= CDR_1
Region 36..49
/label= framework_2
Region 50..65
/label= CDR_2
Region 66..97
/label= framework_3
Region 98..108
/label= CDR_3
Region 109..119
/label= framework_4
WO9640210-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US009847.
07-JUN-1995; 95US-00482982.
15-DEC-1995; 95US-00573289.
(IMCL-) IMCLONE SYSTEMS INC.
(MRCC-) MRC COLLABORATIVE CENT.
Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
WPI; 1997-051897/05.
Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
for inhibiting tumour growth, esp. of late stage prostatic tumour.
Claim 30; Fig 22; 112pp; English.
The present sequence is the heavy chain variable region of the reshaped
human monoclonal antibody (Mab) H225, 225RHE. The Mab is specific for the
human epidermal growth factor (EGF) receptor. The Mab, or a fragment, can
be used to inhibit the growth of tumour cells, especially late stage
prostatic tumour cells in humans, optionally conjugated to a cytotoxic
agent, especially doxorubicin, taxol or cisplatin, or a signal
transduction, ras or cell cycle inhibitor
Sequence 119 AA;
Query Match 100.0%; Score 62; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 98 ALTYDYEFAY 109
RESULT 9
AAW08947
ID AAW08947 standard; protein; 138 AA.
XX

AAW08947;
18-SEP-1997 (first entry)
Heavy chain variable region of 225RA antibody.
Heavy chain; reshaped; monoclonal; antibody; 225RA; human;
epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
late stage; prostatic; prostate; variable region.
Homo sapiens.
WO9640210-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US009847.
07-JUN-1995; 95US-00482982.
15-DEC-1995; 95US-00573289.
(IMCL-) IMCLONE SYSTEMS INC.
(MRCC-) MRC COLLABORATIVE CENT.
Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
WPI; 1997-051897/05.
N-PSDB; AAT49346.
Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
for inhibiting tumour growth, esp. of late stage prostatic tumour.
Claim 31; Fig 20; 112pp; English.
The present sequence is the heavy chain variable region of the reshaped
human monoclonal antibody (Mab) H225, 225RHA. The Mab is specific for the
human epidermal growth factor (EGF) receptor. The Mab, or a fragment, can
be used to inhibit the growth of tumour cells, especially late stage
prostatic tumour cells in humans, optionally conjugated to a cytotoxic
agent, especially doxorubicin, taxol or cisplatin, or a signal
transduction, ras or cell cycle inhibitor
Sequence 138 AA;
Query Match 100.0%; Score 62; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 117 ALTYDYEFAY 127
RESULT 10
AAW08944
ID AAW08944 standard; protein; 138 AA.
XX
AC AAW08944;
XX
DT 18-SEP-1997 (first entry)
XX
DE Heavy chain variable region of C225 antibody.
XX
KW Heavy chain; murine; mouse; monoclonal; antibody; C225; human;
epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
late stage; prostatic; prostate; variable region; chimeric.
Mus; spp.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
FN WO9640210-A1.

```

XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009847.
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (MRCC-) MRC COLLABORATIVE CENT.
XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX DR WPI; 1997-051897/05.
XX DR N-PSDB; AAT49343.
XX CHimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX PS Example IV; Fig 16; 112pp; English.
XX CC The present sequence is the heavy chain variable region of the chimeric
XX CC monoclonal antibody (MAB) C225, which is specific for the human epidermal
XX CC growth factor (EGF) receptor. The MAB, or a fragment, can be used to
XX CC inhibit the growth of tumour cells, especially late stage prostatic
XX CC tumour cells in humans, optionally conjugated to a cytotoxic agent,
XX CC especially doxorubicin, taxol or cisplatin, or a signal transduction, ras
XX CC or cell cycle inhibitor
XX SQ Sequence 138 AA;
XX
XX Query Match 100.0%; Score 62; DB 2; Length 138;
XX Best Local Similarity 100.0%; Pred. No. 0.0065;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ALTYDYEFAY 11
XX DB 117 ALTYDYEFAY 127
XX
XX RESULT 11
XX AAW08942
XX ID AAW08942 standard; protein; 138 AA.
XX AC AAW08942;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 18-SEP-1997 (first entry)
XX XX
XX DE Heavy chain variable region of M225 antibody.
XX XX
XX KW Heavy chain; murine; mouse; monoclonal; antibody; M225; human;
XX KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX KW late stage; prostatic; prostate; variable region.
XX XX
XX OS Mus sp.
XX XX
XX PN WO9640210-A1.
XX XX
XX PD 19-DEC-1996.
XX XX
XX PF 07-JUN-1996; 96WO-US009847.
XX XX
XX PR 07-JUN-1995; 95US-00482982.
XX PR 15-DEC-1995; 95US-00573289.
XX XX
XX PA (IMCL-) IMCLONE SYSTEMS INC.
XX PA (MRCC-) MRC COLLABORATIVE CENT.
XX XX
XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX DR WPI; 1997-051897/05.
XX DR N-PSDB; AAT49340.

```

```

XX CHimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX PS Example IV; Fig 14; 112pp; English.
XX XX
XX CC The present sequence is the heavy chain variable region of the murine
XX CC monoclonal antibody (MAB) M225, which is specific for the human epidermal
XX CC growth factor (EGF) receptor. The MAB, or a fragment, can be used to
XX CC inhibit the growth of tumour cells, especially late stage prostatic
XX CC tumour cells in humans, optionally conjugated to a cytotoxic agent,
XX CC especially doxorubicin, taxol or cisplatin, or a signal transduction, ras
XX CC or cell cycle inhibitor. (Updated on 27-AUG-2003 to correct OS field.)
XX XX
XX SQ Sequence 138 AA;
XX
XX Query Match 100.0%; Score 62; DB 2; Length 138;
XX Best Local Similarity 100.0%; Pred. No. 0.0065;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ALTYDYEFAY 11
XX DB 117 ALTYDYEFAY 127
XX
XX RESULT 12
XX AAW05133
XX ID AAW05133 standard; protein; 240 AA.
XX AC AAW05133;
XX XX
XX DT 29-JAN-1997 (first entry)
XX XX
XX DE Single chain antibody scFv(225).
XX XX
XX KW Single chain antibody; scFv; monoclonal antibody; MAB; EGF;
XX KW epidermal growth factor; receptor; antitumour; cancer; therapy.
XX XX
XX OS Mus; sp.
XX OS Synthetic.
XX OS Chimeric.
XX XX
XX FH Key Location/Qualifiers
XX FT Region 1. .119
XX FT /label= VH_region
XX FT /note= "monoclonal antibody 225 VH"
XX FT Peptide 120. .133
XX FT /label= Linker
XX FT /note= "synthetic spacer peptide"
XX FT Region 134. .240
XX FT /label= VL
XX FT /note= "monoclonal antibody 225 VL"
XX XX
XX PN EP739984-A1.
XX PD 30-OCT-1996.
XX XX
XX XX 26-APR-1995; 95EP-00106275.
XX XX 26-APR-1995; 95EP-00106275.
XX XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX XX
XX PI Wels W, Schmidt M, Groner B;
XX XX
XX DR WPI; 1996-478748/48.
XX DR N-PSDB; AAT42033.
XX XX
XX PT Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX XX
XX PS Example 2; Page 17; 52pp; English.

```

XX scFv(225) (AAW05133) comprises the single-chain binding region of murine
CC monoclonal antibody 225, which is specific for the human epidermal growth
CC factor receptor. It is encoded by plasmid pMW152-225 (see also AAT42033),
CC constructed by cloning MAB 225 VH and VL region cDNAs into plasmid
CC pMW152. Novel bivalent proteins (see also AAW05134-44), some of them
CC including scFv(225) and an effector e.g. cytotoxin, can be produced in
CC bacterial host cells, and are useful as antitumour agents
XX SQ Sequence 240 AA;
Query Match 100.0%; Score 62; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.012; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 98 ALTYDYEFAY 108
RESULT 13
AAW05135
ID AAW05135 standard; protein; 651 AA.
XX AC AAW05135;
XX DT 29-JAN-1997 (first entry)
XX DE scFv(225)-ETA fusion protein.
XX KW Single chain antibody; scFv; monoclonal antibody; MAB; EGF;
XX KW epidermal growth factor; receptor; plasmid pMW202-225; cancer; therapy;
XX KW antitumour; exotoxin A; ETA.
XX OS Mus sp.
XX OS Pseudomonas; aeruginosa.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= Sig_peptide
FT /note= "ompA signal peptide"
FT Peptide 22..38
FT /label= Spacer
FT Protein 39..278
FT /label= scFv(225)
FT Peptide 279..289
FT /label= Spacer
FT Protein 290..651
FT /label= ETA
FT /note= "exotoxin A amino acids 252-613"
XX EP739984-A1.
XX PN 30-OCT-1996.
XX PD 26-APR-1995; 95EP-00106275.
XX PF 26-APR-1995; 95EP-00106275.
XX PR 26-APR-1995; 95EP-00106275.
XX XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX PI Wels W, Schmidt M, Groner B;
XX WPI; 1996-478748/48.
XX DR N-PSDB; AAT42035.
XX XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX XX Example 7; Page 19-20; 52pp; English.

XX scFv(225)-ETA (AAW05135) comprises the single-chain binding region (see
CC also AAW05133) of murine monoclonal antibody 225, which is specific for
CC human epidermal growth factor receptor, joined to exotoxin A (ETA). It is
CC encoded by plasmid pMW202-225 (see also AAT42035) obtd. by ligating an
CC scFv(225) gene (AAT42033) into plasmid pMW200 contg. the Pseudomonas
CC aeruginosa PAK ETA gene. The construct can be used to produce novel
CC bivalent fusion proteins (see also AAW05136-44) in bacterial host cells,
CC for use as antitumour agents
XX SQ Sequence 651 AA;
Query Match 100.0%; Score 62; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYEFAY 11
Db 136 ALTYDYEFAY 146
RESULT 14
AAW05140
ID AAW05140 standard; protein; 892 AA.
XX AC AAW05140;
XX DT 29-JAN-1997 (first entry)
XX DE scFv2(225)/FRP5)-ETA.
XX KW Single chain antibody; scFv; monoclonal antibody; MAB; EGF; etrbB-2;
XX KW epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;
XX KW exotoxin A; ETA; antitumour.
XX OS Mus; sp.
XX OS Pseudomonas; aeruginosa.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= Sig_peptide
FT /note= "ompA signal peptide"
FT Peptide 22..38
FT /label= Spacer
FT Region 39..278
FT /label= scFv(225)
FT Peptide 279..289
FT /label= Spacer
FT Region 290..404
FT /label= ETA
FT /note= "exotoxin A amino acids 252-366"
FT Peptide 405..407
FT /label= Spacer
FT Region 408..647
FT /label= scFv(FRP5)
FT Peptide 648..658
FT /label= Spacer
FT Region 659..892
FT /label= ETA
FT /note= "endotoxin-A amino acids 380-613"
XX EP739984-A1.
XX PN 30-OCT-1996.
XX PD 26-APR-1995; 95EP-00106275.
XX PF 26-APR-1995; 95EP-00106275.
XX PR 26-APR-1995; 95EP-00106275.
XX XX (SANT-) SAN TUMORFORSCHUNGS GMBH.

```

PI Wels W, Schmidt M, Groner B;
XX WPI; 1996-478748/48.
DR N-PSDB; AAT42040.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
XX Example 11; Page 31-33; 52pp; English.
XX
XX scFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region
CC of murine monoclonal antibody 225 (specific for human epidermal growth
CC factor receptor, see also AAW05133) joined to portions of exotoxin A from
CC Pseudomonas aeruginosa and to the single-chain binding region of murine
CC monoclonal antibody FRP5 (specific for human epidermal growth factor
CC receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-
CC 5 (AAT42040). This plasmid can be utilised in the prodn. of the bivalent
CC fusion protein in bacterial (esp. E. coli) host cells. Such fusion
CC proteins (see also AAW05138-44) are useful as antitumour agents
XX
XX Sequence 892 AA;
SQ
Query Match 100.0%; Score 62; DB 2; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYDFAY 11
DB 136 ALTYDYDFAY 146
RESULT 15
AAW05139
ID AAW05139 standard; protein; 892 AA.
XX
XX AAW05139;
XX
XX 29-JAN-1997 (first entry)
DE scFv2(FRP5/225)-ETA (version 1).
XX
XX Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
KW exotoxin A; ETA; antitumour.
XX
XX Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH Peptide 1..21 /label= Sig peptide
FT /note= "ompA signal peptide"
FT Peptide 22..38 /label= Spacer
FT Region 39..278 /label= scFv(FRP5)
FT Peptide 279..289 /label= Spacer
FT Region 290..404 /label= ETA
FT /note= "exotoxin A amino acids 252-366"
FT Peptide 405..407 /label= Spacer
FT Region 408..647 /label= scFv(225)
FT Peptide 648..658 /label= Spacer
FT Region 659..892 /label= ETA
FT /note= "endotoxin-A amino acids 380-613"

```

```

XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX WPI; 1996-478748/48.
XX N-PSDB; AAT42039.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
XX Example 11; Page 28-30; 52pp; English.
XX
XX scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
CC of murine monoclonal antibody FRP5 (specific for human epidermal growth
CC factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
CC A from Pseudomonas aeruginosa and to the single-chain binding region of
CC murine monoclonal antibody 225 (specific for human epidermal growth
CC factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
CC 225 (AAT42039). This plasmid can be utilised in the prodn. of the
CC bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
CC fusion proteins (see also AAW05138-44) are useful as antitumour agents
XX
XX Sequence 892 AA;
SQ
Query Match 100.0%; Score 62; DB 2; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ALTYDYDFAY 11
DB 505 ALTYDYDFAY 515
Search completed: October 6, 2004, 16:29:45
Job time : 57.5789 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 68.5088 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-8
Perfect score: 55
Sequence: 1.RASQSIGTNIH 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```

Database : Published Applications_AA.*
1: /cgm2_6/ptodata/2/pubpaa/US07 PUBCOMB.pcp.*
2: /cgm2_6/ptodata/2/pubpaa/PCT_NEW PUB.pcp.*
3: /cgm2_6/ptodata/2/pubpaa/US06 NEW PUB.pcp.*
4: /cgm2_6/ptodata/2/pubpaa/US06 PUBCOMB.pcp.*
5: /cgm2_6/ptodata/2/pubpaa/US07 NEW PUB.pcp.*
6: /cgm2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pcp.*
7: /cgm2_6/ptodata/2/pubpaa/US08 NEW PUB.pcp.*
8: /cgm2_6/ptodata/2/pubpaa/US08 PUBCOMB.pcp.*
9: /cgm2_6/ptodata/2/pubpaa/US09A PUBCOMB.pcp.*
10: /cgm2_6/ptodata/2/pubpaa/US09B PUBCOMB.pcp
11: /cgm2_6/ptodata/2/pubpaa/US09C PUBCOMB.pcp
12: /cgm2_6/ptodata/2/pubpaa/US09 NEW PUB.pcp.*
13: /cgm2_6/ptodata/2/pubpaa/US10A PUBCOMB.pcp
14: /cgm2_6/ptodata/2/pubpaa/US10B PUBCOMB.pcp
15: /cgm2_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp
16: /cgm2_6/ptodata/2/pubpaa/US10 NEW PUB.pcp.*
17: /cgm2_6/ptodata/2/pubpaa/US60 NEW PUB.pcp.*
18: /cgm2_6/ptodata/2/pubpaa/US60 PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|-----|----|--------------------|-------------------|
| | | Match | \$ | | | |
| 1 | 55 | 100.0 | 11 | 10 | US-09-798-689-32 | Sequence 32, Appl |
| 2 | 55 | 100.0 | 11 | 10 | US-09-996-954B-8 | Sequence 8, Appl |
| 3 | 55 | 100.0 | 11 | 12 | US-10-374-600-100 | Sequence 100, App |
| 4 | 55 | 100.0 | 11 | 15 | US-10-374-531-100 | Sequence 100, App |
| 5 | 55 | 100.0 | 107 | 9 | US-09-991-470-25 | Sequence 25, Appl |
| 6 | 55 | 100.0 | 107 | 12 | US-10-374-600-113 | Sequence 113, App |
| 7 | 55 | 100.0 | 107 | 12 | US-10-374-600-114 | Sequence 114, App |
| 8 | 55 | 100.0 | 107 | 15 | US-10-374-531-113 | Sequence 113, App |
| 9 | 55 | 100.0 | 107 | 15 | US-10-374-531-114 | Sequence 114, App |
| 10 | 55 | 100.0 | 108 | 15 | US-10-412-703A-129 | Sequence 129, App |
| 11 | 55 | 100.0 | 108 | 12 | US-10-374-600-5 | Sequence 5, Appl |
| 12 | 55 | 100.0 | 127 | 12 | US-10-374-600-11 | Sequence 11, Appl |
| 13 | 55 | 100.0 | 127 | 12 | US-10-374-600-15 | Sequence 15, Appl |
| 14 | 55 | 100.0 | 127 | 15 | US-10-374-600-17 | Sequence 17, Appl |
| 15 | 55 | 100.0 | 127 | 15 | US-10-374-531-5 | Sequence 5, Appl |

RESULT 1

```

US-09-798-689-32
/ Sequence 32, Application US/097986899
/ Publication No. US20030103973A1
/ GENERAL INFORMATION:
/ APPLICANT: Rockwell, Patricia
/ APPLICANT: Goldstein, Neil I.
/ TITLE OF INVENTION: Method for Reducing Tumor G
/ TYPE OF INVENTION: Combined With Radiation and
/ FILE REFERENCE: Sequence Listings 1-41 for 381
/ CURRENT APPLICATION NUMBER: US/09/798,689
/ CURRENT FILING DATE: 2001-03-02
/ PRIOR APPLICATION NUMBER: 09/401,163
/ PRIOR FILING DATE: 1999-09-22
/ PRIOR APPLICATION NUMBER: 08/967,113
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 08/706,804
/ PRIOR FILING DATE: 1996-09-03
/ PRIOR APPLICATION NUMBER: 08/476,533
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: 08/326,552
/ PRIOR FILING DATE: 1994-10-20
/ PRIOR APPLICATION NUMBER: 08/196,041
/ PRIOR FILING DATE: 1994-03-10
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 32
/ LENGTH: 11
/ TYPE: PRT
/ ORGANISM: Mouse
US-09-798-689-32

```

ALIGNMENTS

| | |
|---------------|------|
| Sequence 11, | Appl |
| Sequence 15, | Appl |
| Sequence 17, | Appl |
| Sequence 27, | Appl |
| Sequence 4, | Appl |
| Sequence 39, | Appl |
| Sequence 27, | Appl |
| Sequence 27, | Appl |
| Sequence 38, | Appl |
| Sequence 129, | Appl |
| Sequence 130, | Appl |
| Sequence 129, | Appl |
| Sequence 130, | Appl |
| Sequence 49, | Appl |
| Sequence 37, | Appl |
| Sequence 58, | Appl |
| Sequence 39, | Appl |
| Sequence 3, | Appl |
| Sequence 1, | Appl |
| Sequence 1, | Appl |
| Sequence 29, | Appl |
| Sequence 28, | Appl |
| Sequence 9, | Appl |
| Sequence 1, | Appl |
| Sequence 8, | Appl |
| Sequence 48, | Appl |
| Sequence 8, | Appl |
| Sequence 71, | Appl |
| Sequence 30, | Appl |
| Sequence 38, | Appl |

Query Match 100.0%: Score 55; DB 10; Length 11;

| | | | |
|--------------------------|--------|-------------------|-------------------|
| Query Match | 100.0% | Score 33 | Length 11 |
| Best Local Similarity | 100.0% | Pred. No. 0.00059 | |
| Matches 11: Conservative | 0 | Mismatches 0 | Indels 0; Gaps 0; |

Qy 1 RASQSIGTNIH 11


```

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 100:
US-10-374-531-100

Query Match      100.0%; Score 55; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASQSIGTNIH 11
      |||||
Db      1 RASQSIGTNIH 11

RESULT 5
US-09-991-470-25
; Sequence 25, Application US/09991470
; Patent No. US20020173477A1
; GENERAL INFORMATION:
; APPLICANT: Ruy S. Liou
; TITLE OF INVENTION: ANTI-IGE GENE THERAPY
; FILE REFERENCES: 99-5
; CURRENT APPLICATION NUMBER: US/09/991,470
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/397,569
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 107
; TYPE: PRT
; ORGANISM: human/murine
US-09-991-470-25

Query Match      100.0%; Score 55; DB 9; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASQSIGTNIH 11
      |||||
Db      24 RASQSIGTNIH 34

US-09-991-470-25

RESULT 6
US-10-374-600-113
; Sequence 113, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847

US-09-991-470-25

RESULT 7
US-10-374-600-114
; Sequence 114, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
```

```

; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 114:
US-10-374-600-114

Query Match 100.0%; Score 55; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

RESULT 8
US-10-374-531-113
; Sequence 113, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 113:
US-10-374-531-113

Query Match 100.0%; Score 55; DB 15; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

US-10-374-531-113
; Sequence 114, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 114:
US-10-374-531-114

Query Match 100.0%; Score 55; DB 15; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34
```

```
RESULT 10
US-10-412-703A-129
; Sequence 129, Application US/10412703A
; Publication No. US20030219439A1
; GENERAL INFORMATION:
; APPLICANT: Reed et al.
; TITLE OF INVENTION: Recombinant Anti-Interleukin-9 Antibodies
; FILE REFERENCE: IL400US
; CURRENT APPLICATION NUMBER: US/10/412,703A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,728
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/371,683
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 129
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-703A-129

Query Match          100.0%; Score 55; DB 15; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASQSIGTNIH 11
        |||||
Db       24 RASQSIGTNIH 34
        |||||

RESULT 11
US-10-374-600-5
; Sequence 5, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELEPHONE: (212) 425-5288
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-374-600-5
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-374-600-5

Query Match          100.0%; Score 55; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RASQSIGTNIH 11
        |||||
Db       44 RASQSIGTNIH 54
        |||||

RESULT 12
US-10-374-600-11
; Sequence 11, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-374-600-11
```

Query Match 100.0%; Score 55; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 13

US-10-374-600-15
Sequence 15, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09947
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-7200

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-10-374-600-15

Query Match 100.0%; Score 55; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 14

US-10-374-600-17

Sequence 17, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09947
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003

TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-7200

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-10-374-600-17

Query Match 100.0%; Score 55; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 15

US-10-374-531-5

Sequence 5, Application US/10374531
Publication No. US20040006212A1
GENERAL INFORMATION:
APPLICANT: ImClone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York

```

; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <UNKNOWN>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 09/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-374-531-5

```

```

Query Match          100.0%; Score 55; DB 15; Length 127;
Best Local Similarity 100.0%; Pred. NO. 0.0079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

```

```

Search completed: October 6, 2004, 17:09:05
Job time : 70.5088 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:13 ; Search time 36.8596 Seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-8

Perfect score: 55

Sequence: 1 RASQSIGTNIH 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 43 | 78.2 | 127 | 11 Q925S9 | Q925S9 mus musculus |
| 2 | 39 | 70.9 | 108 | 11 Q8VIJ0 | Q8VIJ0 mus musculus |
| 3 | 37 | 67.3 | 212 | 4 Q8NB65 | Q8NB65 homo sapien |
| 4 | 37 | 67.3 | 302 | 17 Q8U0A1 | Q8U0A1 pyrococcus |
| 5 | 37 | 67.3 | 1186 | 11 Q80TU3 | Q80TU3 mus musculus |
| 6 | 36 | 65.5 | 132 | 10 Q8H8X2 | Q8H8X2 oryza sativ |
| 7 | 36 | 65.5 | 239 | 5 Q96SV8 | Q96SV8 caenorhabdi |
| 8 | 36 | 65.5 | 304 | 5 Q7YXU6 | Q7YXU6 caenorhabdi |
| 9 | 36 | 65.5 | 428 | 16 Q8PA89 | Q8PA89 xanthomonas |
| 10 | 36 | 65.5 | 440 | 16 Q8UKJ0 | Q8UKJ0 agrobacteri |
| 11 | 36 | 65.5 | 1434 | 9 Q9AZB3 | Q9AZB3 lactobacill |
| 12 | 35 | 63.6 | 139 | 17 Q8Q0E5 | Q8Q0E5 methanosarc |
| 13 | 35 | 63.6 | 169 | 10 Q38542 | Q38542 clivia mini |
| 14 | 35 | 63.6 | 211 | 2 Q9F8M9 | Q9F8M9 carboxydoch |
| 15 | 35 | 63.6 | 314 | 5 Q8T939 | Q8T939 tribolium c |
| 16 | 35 | 63.6 | 373 | 16 Q7W108 | Q7W108 bordetella |

| | | | | | |
|----|------|------|------|-----------|---------------------|
| 17 | 35 | 63.6 | 373 | 16 Q7W642 | Q7W642 bordetella |
| 18 | 35 | 63.6 | 374 | 16 Q7VVM0 | Q7VVM0 bordetella |
| 19 | 35 | 63.6 | 982 | 10 Q93Y08 | Q93Y08 lycopersico |
| 20 | 35 | 63.6 | 982 | 10 Q65833 | Q65833 lycopersico |
| 21 | 34.5 | 62.7 | 599 | 16 Q7UX08 | Q7UX08 rhodospirell |
| 22 | 34 | 61.8 | 108 | 4 Q9UL83 | Q9UL83 homo sapien |
| 23 | 34 | 61.8 | 194 | 15 Q9HW65 | Q9HW65 pseudomonas |
| 24 | 34 | 61.8 | 302 | 17 Q9VUH2 | Q9VUH2 pyrococcus |
| 25 | 34 | 61.8 | 326 | 16 Q89CL1 | Q89CL1 bradyrhizob |
| 26 | 34 | 61.8 | 326 | 16 Q87PW8 | Q87PW8 vibrio para |
| 27 | 34 | 61.8 | 441 | 11 Q9D749 | Q9D749 mus musculu |
| 28 | 34 | 61.8 | 441 | 11 Q9ERV5 | Q9ERV5 mus musculu |
| 29 | 34 | 61.8 | 502 | 11 Q8BH48 | Q8BH48 mus musculu |
| 30 | 34 | 61.8 | 550 | 3 Q877A3 | Q877A3 aspergillus |
| 31 | 34 | 61.8 | 646 | 11 Q8BSW6 | Q8BSW6 mus musculu |
| 32 | 34 | 61.8 | 699 | 3 Q87616 | Q87616 saccharomyc |
| 33 | 34 | 61.8 | 731 | 8 Q8MA29 | Q8MA29 cyclocheilo |
| 34 | 34 | 61.8 | 788 | 5 Q01500 | Q01500 caenorhabdi |
| 35 | 34 | 61.8 | 806 | 5 Q81711 | Q81711 caenorhabdi |
| 36 | 34 | 61.8 | 898 | 9 Q7Y2C9 | Q7Y2C9 bacterioph |
| 37 | 34 | 61.8 | 918 | 3 Q9HEW7 | Q9HEW7 cladosporiu |
| 38 | 34 | 61.8 | 998 | 10 Q9T0D4 | Q9T0D4 arabidopsis |
| 39 | 34 | 61.8 | 1612 | 16 Q83DF0 | Q83DF0 coxiella bu |
| 40 | 33 | 60.0 | 132 | 16 Q98NC7 | Q98NC7 rhizobium l |
| 41 | 33 | 60.0 | 136 | 16 Q89A73 | Q89A73 buchnera ap |
| 42 | 33 | 60.0 | 152 | 5 Q8MWM1 | Q8MWM1 scarites su |
| 43 | 33 | 60.0 | 167 | 5 Q8MX00 | Q8MX00 dicaelus am |
| 44 | 33 | 60.0 | 170 | 16 Q88ST3 | Q88ST3 lactobacill |
| 45 | 33 | 60.0 | 224 | 16 Q98M72 | Q98M72 rhizobium l |

ALIGNMENTS

RESULT 1

Q925S9 PRELIMINARY; PRT; 127 AA.

ID Q925S9
AC Q925S9;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Immunoglobulin light chain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RA MEDLINE=99306687; PubMed=10380019;
RA Tripkathi P.K., Qin H., Bhattacharya-Chatterjee M., Ceriani R.L.,
RA Foon K.A., Chatterjee S.K.;
RT "Construction and characterization of a chimeric fusion protein
RT consisting of an anti-idiotypic antibody mimicking a breast cancer-
RT associated antigen and the cytokine GM-CSF";
RL Hybridoma 18193-202(1999).
DR EMBL; AF124721; AAKS5120.1; -.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
FT NON_TER 127
SQ SEQUENCE 127 AA; 13794 MW; 13F61BEBBB981FA5 CRC64;

Query Match 78.2%; Score 43; DB 11; Length 127;

Best Local Similarity 72.7%; Pred. No. 0.9;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RASQSIGTNIH 11

Db 44 RASQDIGINLH 54

```

RESULT 2
Q8VJ0 PRELIMINARY; PRT; 108 AA.
AC Q8VJ0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Anti-DNA light chain (Fragment).
GN VK19.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C3H/HeJ-lpr/lpr;
RX MEDLINE=96409289; PubMed=8814271;
RA Wloch M.K., Alexander A.L., Phippen A.M., Pisetsky D.S., Gilkeson G.S.;
RT "Differences in V kappa gene utilization and VH CDR3 sequence among
anti-DNA from C3H-lpr/lpr mice and lupus mice with nephritis.";
RL Eur. J. Immunol. 26:2225-2233(1996).
DR EMBL; U59155; AAB02917.1; -.
DR FIC; A33933; A33933.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0046; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11859 MW; 68506D75613DBFEE CRC64;

Query Match 70.9%; Score 39; DB 11; Length 108;
Best Local Similarity 60.0%; Pred. No. 4.9;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
Db 24 KASQNVGTNV 33

RESULT 3
Q8NB65 PRELIMINARY; PRT; 212 AA.
AC Q8NB65;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ34176.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Ishibashi T., Kaneshori K., Yoshida M., Watanabe S., Ishida S., Ono Y.,
RA Hotta T., Hirakawa S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK091495; BAC03676.1; -.
DR GO; GO:0003676; F: nucleic acid binding; IEA.
DR GO; GO:0016032; P: viral life cycle; IEA.
DR InterPro; IPR000721; Gag_P24.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR001878; Znf_CCHC.

```

```

DR Pfam; PF00607; Gag_P24; 1.
DR Pfam; PF00098; Zf_CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PSS0158; Zf_CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 212 AA; 23378 MW; 2070DE43BD5EE228 CRC64;

Query Match 67.3%; Score 37; DB 4; Length 212;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 97 RACQDVGTGLH 107

RESULT 4
Q8U0A1 PRELIMINARY; PRT; 302 AA.
AC Q8U0A1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative ribose ABC transporter.
GN PF1698.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010267; AL81822.1; -.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 302 AA; 32345 MW; DDE312589EB58EA CRC64;

Query Match 67.3%; Score 37; DB 17; Length 302;
Best Local Similarity 45.5%; Pred. No. 40;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 169 RAAEAMGVNVH 179

RESULT 5
Q80TU3 PRELIMINARY; PRT; 1186 AA.
AC Q80TU3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MKTAA0684 protein (Fragment).
GN MKTAA0684.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuro R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara C., Koga H.;

```

```

RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT 11. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122345; BAC65627.1; -.
DR InterPro; IPR003613; Znf_modRING.
DR Pfam; PF04564; U-box; 1.
DR SMART; SMO0504; Ubox; 1.
FT NON_TER 1
SQ SEQUENCE 1186 AA; 134813 MW; 2FC25DC03820A175 CRC64;

Query Match 67.3%; Score 37; DB 11; Length 1186;
Best Local Similarity 66.7%; Pred.No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 SQSIGTNIH 11
|||:|:|
DB 63 SQSLGLNVH 71

RESULT 6
Q8H9X2 PRELIMINARY; PRT; 132 AA.
AC Q8H9X2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSUNBA0062C05.28.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUNBA0062C05 genomic sequence.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084023; AAN05573.1; -.
DR EMBL; AB017110; AAP54523.1; -.
KW Hypothetical protein.
SQ SEQUENCE 132 AA; 14290 MW; 09412E1E292D2D9E CRC64;

Query Match 65.5%; Score 36; DB 10; Length 132;
Best Local Similarity 85.7%; Pred.No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 SGTNIH 11
||||:|
DB 94 SGTNVH 100

RESULT 7

```

```

Q965V8 PRELIMINARY; PRT; 299 AA.
ID Q965V8;
AC Q965V8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Y43B11AL.2.
GN Y43B11AL.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Cordes M., Maupin R.;
RT "The sequence of C. elegans cosmid Y43B11AL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024779; AAK68485.1; -.
DR WormPep; Y43B11AL.2; CE28264.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 299 AA; 34315 MW; 11C37C5C63A4D70 CRC64;

Query Match 65.5%; Score 36; DB 5; Length 299;
Best Local Similarity 80.0%; Pred.No. 63;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASQSIGTNIH 11
|||||
DB 25 ASQSIGDAIH 34

RESULT 8
Q7YXU6 PRELIMINARY; PRT; 304 AA.
ID Q7YXU6;
AC Q7YXU6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Y43B11AL.2.
GN Y43B11AL.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN=Bristol N2;
 RA Cordes M., Maupin R.;
 RT "The sequence of C. elegans cosmid Y43B11AL.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024779; AAP86615.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 304 AA; 34724 MW; 7F7F6C54C4F69589 CRC64;

Query Match 65.5%; Score 36; DB 5; Length 304;
 Best Local Similarity 80.0%; Pred. No. 64;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASQSIGTNIH 11
 |||||
 Db 36 ASQSIGDAIH 45

RESULT 9

Q8PA89 PRELIMINARY; PRT; 428 AA.
 ID Q8PA89;
 AC Q8PA89;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hydroxylase.
 GN MBTG OR XC1598.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=2022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., de Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Canavaral F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,

RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE012261; AAM40892.1; -
 KW Complete proteome.
 SQ SEQUENCE 428 AA; 47012 MW; FA7C59345838FEE9 CRC64;

Query Match 65.5%; Score 36; DB 16; Length 428;
 Best Local Similarity 63.6%; Pred. No. 94;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 75 RARQKIGVSIH 85

RESULT 10

Q8UKJ0 PRELIMINARY; PRT; 440 AA.
 ID Q8UKJ0;
 AC Q8UKJ0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein Atu5112.
 GN ATU5112 OR AGR_PAT 166.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley P., Tirney S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008936; AAL45802.1; -
 DR EMBL; AE007884; AAK90488.1; -
 DR PIR; AD3173; AD3173.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 KW Hypothetical protein; Plasmid; Complete proteome.
 SQ SEQUENCE 440 AA; 48826 MW; 583770DD1576b6f8 CRC64;

Query Match 65.5%; Score 36; DB 16; Length 440;
 Best Local Similarity 54.5%; Pred. No. 97;

```

Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1 RASQSIGTNIH 11
Db      229 RIAQEIQLNVH 239

RESULT 11
Q9AZB3      PRELIMINARY;      PRT;      1434 AA.
ID Q9AZB3
AC Q9AZB3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative minor tail protein.
GN ORF1434.
OS Lactobacillus johnsonii prophage Lj965.
OC Viruses
OX NCBI_TaxID=139870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20455575; PubMed=10998330;
RA Desiere F., Pridmore R.D., Brussow H.;
RT "Comparative genomics of the late gene cluster from lactobacillus
RT phages.", 275:294-305(2000).
RL Virology
DR EMBL; AF195900; AAK27908.1; -
SQ SEQUENCE 1434 AA; 158446 MW; 0BE8B85A963BC42C CRC64;

Query Match      65.5%; Score 36; DB 9; Length 1434;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches      6;  Conservative      3;  Mismatches      0;  Indels      0;  Gaps      0;

QY      3 SQSIGTNIH 11
Db      1339 AQNIGSNH 1347

RESULT 12
Q8Q0E5      PRELIMINARY;      PRT;      139 AA.
ID Q8Q0E5
AC Q8Q0E5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein MW0192.
GN MW0192.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCX 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Battacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013243; AM29888.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 139 AA; 16170 MW; 28AF39CBFFCD9752 CRC64;

Query Match      63.6%; Score 35; DB 17; Length 139;
Best Local Similarity 54.5%; Pred. No. 43;
Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1 RASQSIGTNIH 11
Db      1 RASQSIGTNIH 11

```

```

Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1 RASQSIGTNIH 11
Db      229 RIAQEIQLNVH 239

RESULT 11
Q9AZB3      PRELIMINARY;      PRT;      1434 AA.
ID Q9AZB3
AC Q9AZB3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative minor tail protein.
GN ORF1434.
OS Lactobacillus johnsonii prophage Lj965.
OC Viruses
OX NCBI_TaxID=139870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20455575; PubMed=10998330;
RA Desiere F., Pridmore R.D., Brussow H.;
RT "Comparative genomics of the late gene cluster from lactobacillus
RT phages.", 275:294-305(2000).
RL Virology
DR EMBL; AF195900; AAK27908.1; -
SQ SEQUENCE 1434 AA; 158446 MW; 0BE8B85A963BC42C CRC64;

Query Match      65.5%; Score 36; DB 9; Length 1434;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches      6;  Conservative      3;  Mismatches      0;  Indels      0;  Gaps      0;

QY      3 SQSIGTNIH 11
Db      1339 AQNIGSNH 1347

RESULT 12
Q8Q0E5      PRELIMINARY;      PRT;      139 AA.
ID Q8Q0E5
AC Q8Q0E5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein MW0192.
GN MW0192.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCX 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Battacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013243; AM29888.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 139 AA; 16170 MW; 28AF39CBFFCD9752 CRC64;

Query Match      63.6%; Score 35; DB 17; Length 139;
Best Local Similarity 54.5%; Pred. No. 43;
Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

QY      1 RASQSIGTNIH 11
Db      1 RASQSIGTNIH 11

```

```

Db      90 KGSDSIGRNLH 100

RESULT 13
Q39542      PRELIMINARY;      PRT;      169 AA.
ID Q39542
AC Q39542
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lectin precursor (Fragment).
OS Clivia miniata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Clivia.
OX NCBI_TaxID=16049;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RX MEDLINE=94250846; PubMed=8193308;
RA Van Damme E.J., Smeets K., Van Leuven F., Peumans W.J.;
RT "Molecular cloning of mannose-binding lectins from Clivia miniata.";
RL Plant Mol. Biol. 24:825-830(1994).
DR EMBL; L16512; AAA19911.1; -
DR PIR; S43782; S43762.
DR HSP; P30617; LQPC.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR001480; B_lectin.
DR Pfam; PF01453; AGglutinin; 1.
DR SMART; SM00108; B_lectin; 1.
KW Signal.
FT NON_TER      1      1
FT SIGNAL      <1      30      POTENTIAL.
FT CHAIN      31      169      LECTIN.
SQ SEQUENCE 169 AA; 18138 MW; 855FDBA3FF7F0B5A CRC64;

Query Match      63.6%; Score 35; DB 10; Length 169;
Best Local Similarity 70.0%; Pred. No. 53;
Matches      7;  Conservative      1;  Mismatches      2;  Indels      0;  Gaps      0;

QY      2 ASQSIGTNIH 11
Db      130 ASWATGTNIH 139

RESULT 14
Q9F8M9      PRELIMINARY;      PRT;      211 AA.
ID Q9F8M9
AC Q9F8M9
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 4-hydroxybutyrate CoA transferase (Fragment).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae;
OC Carboxydothermus.
OX NCBI_TaxID=129958;
RN [1]
RP SEQUENCE FROM N.A.
RA Gonzalez J.M., Robb F.T.;
RT "A genomic survey of the extreme thermophilic, CO-utilizing bacterium
RT Carboxydothermus hydrogenoformans.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF244604; AAG23553.1; -
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006084; P:acetyl-CoA metabolism; IEA.
DR InterPro; IPR003702; ActCoA_hydro.
DR Pfam; PF02550; AcetylCoA_hydro; 1.
KW Transferase.
FT NON_TER      1      1
FT NON_TER      211      211
SQ SEQUENCE 211 AA; 23149 MW; C3038601E0C79E82 CRC64;

```

Query Match 63.6%; Score 35; DB 2; Length 211;
 Best Local Similarity 70.0%; Pred. No. 68;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ASQSIGTNIH 11
 :||:|||||
 Db 183 ASESIGTKIY 192

RESULT 15

QST939
 ID Q8T939 PRELIMINARY; PRT; 314 AA.
 AC Q8T939;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE ultrabithorax.
 GN UBX.
 OS Tribolium castaneum (Red flour beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Tenebrionidae; Tribolium.
 OX NCBI_TaxID=7070;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Galant R., Carroll S.;
 RT "Evolution of a transcriptional repression domain in an insect Hox
 protein.";
 RL Nature 0:0-0(2002).
 CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AY074761; AAL71874.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEIDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 314 AA; 33942 MW; 7679AD11EA107B48 CRC64;

Query Match 63.6%; Score 35; DB 5; Length 314;
 Best Local Similarity 60.0%; Pred. No. 11e-02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASQSIGTNIH 11
 :||:|||||
 Db 189 SSQPVGQIH 198

Search completed: October 6, 2004, 16:33:57
 Job time : 41.8596 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 6.75439 Seconds
(without alignments)
84.800 Million cell updates/sec

Title: US-09-635-974A-8
Perfect score: 55
Sequence: 1 RASQSIGTNIH 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 50 | 90.9 | 115 | 1 KV51_MOUSE | P01642 mus musculus |
| 2 | 37 | 67.3 | 1173 | 1 UB4B_MOUSE | Q98800 mus musculus |
| 3 | 37 | 67.3 | 1302 | 1 UB4B_HUMAN | Q95155 homo sapien |
| 4 | 36 | 65.5 | 886 | 1 YE85_SCHPO | O14302 schizosacch |
| 5 | 35 | 63.6 | 261 | 1 CATG_MOUSE | P28293 mus musculus |
| 6 | 35 | 63.6 | 432 | 1 RMUC_RICCN | Q92975 rickettsia |
| 7 | 34 | 61.8 | 128 | 1 KV3K_HUMAN | P06311 homo sapien |
| 8 | 34 | 61.8 | 252 | 1 TRFP_DROME | P91641 drosophila |
| 9 | 34 | 61.8 | 368 | 1 YB11_YEAST | P38180 saccharomyc |
| 10 | 34 | 61.8 | 441 | 1 DGT1_SHEON | Q8ee42 shewanella |
| 11 | 34 | 61.8 | 536 | 1 ARP_ARATH | P45951 arabidopsis |
| 12 | 34 | 61.8 | 1454 | 1 CP5A_CAEEL | Q944c2 caenorhadi |
| 13 | 34 | 61.8 | 3164 | 1 TEGU_HSV11 | P10220 herpes simp |
| 14 | 33 | 60.0 | 87 | 1 YJHC_LACLA | Q9cgz6 lactococcus |
| 15 | 33 | 60.0 | 129 | 1 KV3H_HUMAN | P04207 homo sapien |
| 16 | 33 | 60.0 | 130 | 1 KV5G_MOUSE | P01639 mus musculus |
| 17 | 33 | 60.0 | 318 | 1 ATPC_YEAST | P32453 saccharomyc |
| 18 | 33 | 60.0 | 336 | 1 YTCB_PSESO | P27103 pseudomonas |
| 19 | 33 | 60.0 | 337 | 1 MEUI_YEAST | Q07938 saccharomyc |
| 20 | 33 | 60.0 | 386 | 1 YAG6_VETJA | Q58466 mechanococc |
| 21 | 33 | 60.0 | 416 | 1 PUR2_CMAJE | Q9pn47 campylobact |
| 22 | 33 | 60.0 | 444 | 1 NIFN_ANASP | Q44145 anabaena sp |
| 23 | 33 | 60.0 | 1520 | 1 ACDF_ECOLI | Q46837 escherichia |
| 24 | 32 | 58.2 | 132 | 1 SP42_AERPE | Q9vg22 aeropyrum p |
| 25 | 32 | 58.2 | 274 | 1 DAPD_ACTPL | P41396 actinobacil |
| 26 | 32 | 58.2 | 293 | 1 YNEJ_ECOLI | P77309 escherichia |
| 27 | 32 | 58.2 | 329 | 1 YCLC_PSEPU | Q47100 pseudomonas |
| 28 | 32 | 58.2 | 393 | 1 VASS_BPM52 | P03610 bacterioph |
| 29 | 32 | 58.2 | 455 | 1 VL2_HPV11 | P04013 human papil |
| 30 | 32 | 58.2 | 468 | 1 YPS7_CAEEL | Q20085 caenorhadi |
| 31 | 32 | 58.2 | 473 | 1 VL2_HPV16 | P03107 human papil |
| 32 | 32 | 58.2 | 474 | 1 PNT5_HAEIN | P43010 haemophilus |
| 33 | 32 | 58.2 | 483 | 1 GLG3_SOLTU | P55243 solanum tub |

| | | | | | |
|----|----|------|------|--------------|--------------------|
| 34 | 32 | 58.2 | 595 | 1 YE85_SCHPO | O14301 schizosacch |
| 35 | 32 | 58.2 | 637 | 1 NUSM_STRPU | P15552 strongyloce |
| 36 | 32 | 58.2 | 657 | 1 HCY_PALVU | P80888 palinurus v |
| 37 | 32 | 58.2 | 727 | 1 MUTB_PROFR | P11653 propionibac |
| 38 | 32 | 58.2 | 972 | 1 ORC4_SCHPO | Q9y794 schizosacch |
| 39 | 32 | 58.2 | 1141 | 1 CN3A_HUMAN | Q14432 homo sapien |
| 40 | 32 | 58.2 | 1520 | 1 ACDF_VIECH | Q9kt04 vibrio chol |
| 41 | 31 | 56.4 | 61 | 1 XYL8_ACTMI | P12867 actinoplan |
| 42 | 31 | 56.4 | 89 | 1 YZFA_ECOLI | P08339 escherichia |
| 43 | 31 | 56.4 | 103 | 1 CHLB_SELMO | P17856 selaginella |
| 44 | 31 | 56.4 | 108 | 1 KV1F_HUMAN | P01598 homo sapien |
| 45 | 31 | 56.4 | 108 | 1 KV3V_MOUSE | P01674 mus musculu |

ALIGNMENTS

RESULT 1

KV51_MOUSE

ID KV51_MOUSE STANDARD; PRT; 115 AA.

AC P01642;

DT 21-JUL-1986 (Rel. 01, Last Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Ig kappa chain V-V region L7 precursor (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=81220975; PubMed=6264318;

RA Pech M., Hochtl J., Schnell H., Zachau H.G.;

RT "Differences between germ-line and rearranged immunoglobulin V kappa

coding sequences suggest a localized mutation mechanism.";

RL Nature 291:668-670(1981).

CC -!- MISCELLANEOUS: THERE APPEAR TO BE TWO POSSIBLE SPLICE JUNCTIONS AT

THE 3' END OF THE INTRON. THE ALTERNATE WOULD CODE FOR A PROTEIN

LACKING RESIDUES 17-19.

DR PIR; A01925; KMSL7.

DR PDB; 1J10; 18-FEB-03.

DR PDB; 1J1P; 18-FEB-03.

DR PDB; 1J1X; 18-FEB-03.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PS00835; IG-LIKE; 1.

KW Immunoglobulin V region; Signal; 3D-structure.

FT SIGNAL 1 20

FT CHAIN 1 >115 IG KAPPA CHAIN V-V REGION L7.

FT DOMAIN 21 43 FRAMEWORK-1.

FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.

FT DOMAIN 55 69 FRAMEWORK-2.

FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.

FT DOMAIN 77 108 FRAMEWORK-3.

FT DOMAIN 109 >115 COMPLEMENTARITY-DETERMINING-3.

FT DISULFID 43 108 BY SIMILARITY.

FT NON_TER 115 115

SQ SEQUENCE 115 AA; 12615 MW; C17BEC758C577E00 CRC64;

Query Match 90.9%; Score 50; DB 1; Length 115;

Best Local Similarity 90.9%; Pred. No. 0.0034;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11

DB 44 RASQSIGTSIH 54

RESULT 2

UB4B_MOUSE

ID UB4B_MOUSE STANDARD; PRT; 1173 AA.


```
Query Match      67.3%; Score 37; DB 1; Length 1302;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

QY 3 SOSIGTNIH 11
 Db 50 SOSLGNVH 58

RESULT 4
 YE86 SCHPO
 ID YE86 SCHPO STANDARD; PRT; 886 AA.
 AC O14302;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C9G1.06c in chromosome I.
 GN SPAC9G1.06C.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Woestl D., Hilbert H.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe.";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -1- SIMILARITY: TO YEAST YDL17W.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z98763; CAB11490.1; -
 CC PIR; T39229; T39229.
 CC HSP; P29355; 1SEP.
 CC GeneDB SPombe; SPAC9G1.06c; -
 CC InterPro; IPR001452; SH3
 CC InterPro; IPR002931; Trnsglucase_like.
 CC Pfam; PF00018; SH3; 1.
 CC ProDom; PD000066; SH3; 1.
 CC SMART; SM00326; SH3; 1.
 CC SMART; SM00460; TGC; 1.
 CC PROSITE; PS00002; SH3; 1.
 CC Hypothetical protein; SH3 domain.
 KW PROSITE; PS00002; SH3; 1.
 KW Hypothetical protein; SH3 domain.
 FT DOMAIN 6 57 SH3.

FT DOMAIN 607 610 POLY-SER.
 SQ SEQUENCE 886 AA; 98262 MW; D0C4879C1882E869 CRC64;
 Query Match 65.5%; Score 36; DB 1; Length 886;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 SIGTNIH 11
 Db 451 SVGTNIH 457

RESULT 5
 CATG MOUSE
 ID CATG MOUSE STANDARD; PRT; 261 AA.
 AC P28293;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cathepsin G precursor (EC 3.4.21.20) (Vimentin-specific protease)
 DE (VSP).
 DE CTSG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Connective tissue;
 RX MEDLINE=93200524; PubMed=8453108;
 RA Heusel J.W., Scarpati E.M., Jenkins N.A., Gilbert D.J., Copeland N.G.,
 RA Shapiro S.D., Ley T.J.;
 RA "Molecular cloning, chromosomal location, and tissue-specific
 RA expression of the murine cathepsin G gene.";
 RL Blood 81:1614-1623(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kulmburg P., Baumrucker T., Werner F.;
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Leadon X A1;
 RA Huang R., Aveskogh M., Hellman L.T.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 21-60.
 RX MEDLINE=92249339; PubMed=1577012;
 RA Nakamura N., Tsuru A., Hirayoshi K., Nagata K.;
 RL "Purification and characterization of a vimentin-specific protease in
 RL mouse myeloid leukemia cells. Regulation during differentiation and
 RL identity with cathepsin G.";
 RL Eur. J. Biochem. 205:947-954(1992).
 CC -1- FUNCTION: This vimentin-specific protease may regulate the
 CC reorganization of vimentin filaments, occurring during cell
 CC differentiation, movement and mitosis.
 CC -1- CATALYTIC ACTIVITY: Specificity similar to chymotrypsin C.
 CC -1- SUBCELLULAR LOCATION: STRONGLY ASSOCIATED WITH MEMBRANES.
 CC -1- SIMILARITY: Belongs to peptidase family S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M36801; AAA37376.1; -
 CC EMBL; X70057; CAA49661.1; -
 CC EMBL; X78544; CAA55290.1; -
 CC PIR; S40162; S40162.
 CC HSP; P08311; 1CGH.
 CC MEROPS; S01.133; -.

```
DR MGD; MGI:88563; Cysg.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM0020; TRYP SP; 1.
DR PROSITE; PS00240; TRYP SIN DOM; 1.
DR PROSITE; PS00134; TRYP SIN HIS; 1.
DR PROSITE; PS00135; TRYP SIN SER; 1.
KW Hydrolase; Serine protease; Zymogen; Glycoprotein; Signal;
KW Intermediate filament; Membrane.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 20 ACTIVATION PEPTIDE.
FT CHAIN 21 261 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 64 64 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 201 201 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 65 BY SIMILARITY.
FT DISULFID 142 207 BY SIMILARITY.
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 51 51 G -> S (IN REF. 4).
FT CONFLICT 56 56 E -> G (IN REF. 4).
FT CONFLICT 60 60 L -> P (IN REF. 4).
FT SEQUENCE 261 AA; 29095 MW; 5EFA1A6E10E1D7FC CRC64;
Query Match 63.6%; Score 35; DB 1; Length 261;
Best Local Similarity 70.0%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
DB 149 RVSQSGTNI 158

RESULT 6
RMUC_RICCN STANDARD; PRT; 432 AA.
ID RMUC_RICCN 29 COILED COIL (POTENTIAL).
AC Q92GR5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA recombination protein rmuc homolog.
GN RMUC OR RC1057.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wallish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RL "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
Science 293:2093-2098(2001).
CC -!- FUNCTION: Involved in DNA recombination (By similarity).
CC -!- SIMILARITY: Belongs to the rmuc family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; Z00021; CAA77316.1; -
DR PIR; A01899; K3HU41.
DR HSP; P01607; IREI.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SMO0406; igv; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 128 IG KAPPA CHAIN V-III REGION IARC/BL41.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 117 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 118 128 JK1 SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14070 MW; CC8957F0FE3B9012 CRC64;
Query Match 61.8%; Score 34; DB 1; Length 128;
Best Local Similarity 60.0%; Pred. No. 7.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
DB 44 RASQSVSNL 53

KW DNA recombination; Coiled coil; Complete proteome.
FT DOMAIN 29 88 COILED COIL (POTENTIAL).
SQ SEQUENCE 432 AA; 48721 MW; 73DF45A996A2F652 CRC64;
Query Match 63.6%; Score 35; DB 1; Length 432;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SQSIGTNIH 11
DB 357 SQKIGNNLH 365

RESULT 7
KV3K_HUMAN STANDARD; PRT; 128 AA.
ID KV3K_HUMAN 29 COILED COIL (POTENTIAL).
AC P06311;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG kappa chain V-III region IARC/BL41 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041852; PubMed=2997711;
RA Klobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III.";
RL Nucleic Acids Res. 13:6499-6513(1985).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; Z00021; CAA77316.1; -
DR PIR; A01899; K3HU41.
DR HSP; P01607; IREI.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SMO0406; igv; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 128 IG KAPPA CHAIN V-III REGION IARC/BL41.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 117 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 118 128 JK1 SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14070 MW; CC8957F0FE3B9012 CRC64;
Query Match 61.8%; Score 34; DB 1; Length 128;
Best Local Similarity 60.0%; Pred. No. 7.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
DB 44 RASQSVSNL 53
```

```

RESULT 8
TRFP_DROME
ID TRFP_DROME STANDARD; PRT; 252 AA.
AC P91641; Q9VLT4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trifunctional protein (transcription mediator-related protein).
GN TRFP OR TMR OR CGI8780/CG18267.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
SEQUENCE FROM N.A.
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Crowley T.E.;
RT "Mutations near the Trf cluster cause a premeiotic defect in the
RL Drosophila male germ line.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bertone P., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshtina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Paclele J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RP REVISIONS.
RX MEDLINE=22426069; PubMed=125317572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celnik S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

```

```

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.W.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R.A., Gonzalez M., Guarin H., Krommiller B., Li P.W., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclele J.M., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K.H., Yu C., Lewis S.E., Rubin G.M.,
RA Celnik S.E.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May regulate transcription of class II genes through
CC association with the RNA polymerase II-SRB complex (By
CC similarity).
CC -!- SUBUNIT: Component of an RNA polymerase II-SRB complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y10975; CAA11871.1; --
CC EMBL; AE003619; AAP52599.2; --
CC EMBL; BT003473; AAO39476.1; --
CC FlyBase; FBgn0013531; Trfp.
CC Transcription regulation; Nuclear protein.
CC CONFLICT 220 220 H -> Q (IN REF. 1).
CC SEQUENCE 252 AA; 27862 MW; 508AB5A43DAC52F5 CRC64;
DR
DR Query Match 61.8%; Score 34; DB 1; Length 252;
DR Best Local Similarity 55.6%; Pred. No. 15;
DR Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 SQSIGNIH 11
Db 233 SQVGVNVH 241
ID YB11_YEAST STANDARD; PRT; 368 AA.
AC P38130;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 40.8 kDa protein in RHK1-PET112 intergenic region.
GN YBL081W OR YBL0722.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288C;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 11:1103-1112(1995).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=95274327; PubMed=7754714;
RA Cusick M.E.;
RT "Sequence of a segment of yeast chromosome II shows two novel genes,

```

```

RT one almost entirely hydrophobic and the other extremely
RL asparagine-serine rich."
RL Yeast 10.1251-1256(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X79489; CAA56026.1; -
DR EMBL; Z35842; CAA84902.1; -
DR EMBL; M89908; AAA75353.1; -
DR PIR; S46601; S46601.
DR GerOnline; 138516; -.
DR SGD; S0000177; YBLO81W.
KW Hypothetical protein.
SQ SEQUENCE 368 AA; 40767 MW; D52EAA7AD85CSD14 CRC64;
Query Match 61.8%; Score 34; DB 1; Length 368;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 SOSIGNIH 11
DB 294 SSSIGNIN 302
| | | | |
| | | | |
RESULT 10
DGT1-SHEON
ID DGT1-SHEON STANDARD; PRT; 441 AA.
AC Q8EEA2;
RX MEDLINE=22297686; PubMed=12368813;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Deoxyguanosine triphosphate triphosphohydrolase-like protein.
GN S02485.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis".
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- SIMILARITY: Belongs to the dGTPase family. Subfamily 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF015690; AAN55516.1; -.
DR TIGR; S02485; -.
DR HAMAP; MF_01212; -; 1.
DR InterPro; IPR006261; dGTP_tripase.
DR InterPro; IPR006674; HD.

```

```

DR InterPro; IPR003607; Met_phosphohydro.
DR Pfam; PF01965; HD; 1.
DR SMART; SM00471; HDC; 1.
DR TIGRFAMs; TIGR01353; dGTP_tripase; 1.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 441 AA; 50619 MW; C1DA17F434ED9514 CRC64;
Query Match 61.8%; Score 34; DB 1; Length 441;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 RASQSIGTNIH 11
DB 399 RASQQLGLNSH 409
| | | | |
| | | | |
RESULT 11
ARP_ARATH
ID ARP_ARATH STANDARD; PRT; 536 AA.
AC P45951;
RX STRAIN=cv. Columbia;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Apurinic endonuclease-redox protein (DNA-(apurinic or apyrimidinic
DE site) lyase) (EC 4.2.99.18).
GN ARP OR REF OR AT2G41460 OR T26J13.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RA thaliana."
RT Nature 402:761-768(1999).
RN [2]
RP SEQUENCE OF 10-536 FROM N.A.
RX STRAIN=cv. Columbia; TISSUE=Callus;
RX MEDLINE=94211851; PubMed=7512729;
RA Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT "The Arabidopsis thaliana apurinic endonuclease Arp reduces human
RT transcription factors Fos and Jun."
RL Proc. Natl. Acad. Sci. U.S.A. 91:3299-3303(1994).
CC -!- FUNCTION: Repairs oxidative DNA damages, seems also to act as a
CC redox factor. Is multifunctional and may be involved both in DNA
CC repair and in the regulation of transcription.
CC -!- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or
CC apyrimidinic site in DNA is broken by a beta-elimination reaction,
CC leaving a 3'-terminal unsaturated sugar and a product with a
CC terminal 5'-phosphate.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in the siliques, flowers, and stems.
CC -!- A high level expression is seen in the leaves.
CC -!- SIMILARITY: Belongs to the DNA repair enzymes AP/exoA family.
CC -!- SIMILARITY: Contains 1 SAP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AC004625; AAC23731.1; -.
CC EMBL; X76912; CAA54234.1; -.
CC PIR; T02441; T02441.
CC HSP; P27695; LE9N.
CC InterPro; IPR000097; Apendonclsel.
CC InterPro; IPR005135; Exo_endo_phos.
CC InterPro; IPR004808; ExoIII_xth.
CC InterPro; IPR003034; SAP.
CC Pfam; PF03372; Exo_endo_phos; 1.
CC Pfam; PF02037; SAP; 1.
CC SMART; SM00513; SAP; 1.
CC TIGRfams; TIGR00195; exodNase_III; 1.
CC TIGRfams; TIGR00633; xth; 1.
CC PROSITE; PS00726; AP_NUCLEASE_F1_1; 1.
CC PROSITE; PS00727; AP_NUCLEASE_F1_2; 1.
CC PROSITE; PS00728; AP_NUCLEASE_F1_3; 1.
CC PROSITE; PS0800; SAP; 1.
CC DNA repair; Lyase; Nuclear protein.
KW DOMAIN 1 278 HIGHLY CHARGED; INCREASES THE AFFINITY OF
FT ARP FOR DNA.
FT DOMAIN 97 131
FT DOMAIN 279 536 AP ENDONUCLEASE.
FT METAL 313 536 MAGNESIUM OR MANGANESE (BY SIMILARITY).
FT ACT_SITE 527 537 GENERAL BASE (BY SIMILARITY).
FT SEQUENCE 536 AA; 60260 MW; 5C1FC17EA991D27B CRC64;

Query Match 61.8%; Score 34; DB 1; Length 536;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SOSIGTNIH 11
DB 506 SOSIAANH 514

RESULT 12
CPSA CAEEL STANDARD; PRT; 1454 AA.
AC Q9NAC2; (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable cleavage and polyadenylation specificity factor, 160 kDa
DE subunit (CPSF 160 kDa subunit).
GN Y76B12C.7
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1] SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RC Cordes M.;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [12]
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CPSF plays a key role in pre-mRNA 3'-end formation,
CC recognizing the AAUAAA signal sequence and interacting with
CC poly(A) polymerase and other factors to bring about cleavage and
CC poly(A) addition. This subunit is involved in the RNA recognition
CC step of the polyadenylation reaction (By similarity).
CC -!- SUBUNIT: CPSF is a heterotrimer composed of four distinct
CC subunits 160, 100, 70 and 30 kDa (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the CPSF160 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AC024211; AAF36067.2; -.
CC WormPep; Y76B12C.7; CE29932.
CC GO; GO:005847; C-mRNA cleavage and polyadenylation specifici. .; NAS.
CC GO; GO:0030364; F:cleavage/polyadenylation specificity factor. .; NAS.
CC GO; GO:0006378; P:mRNA polyadenylation; NAS.
CC InterPro; IPR004871; CPSF_A_C.
CC Pfam; PF03178; CPSF_A; 1.
CC mRNA processing; Nuclease; RNA-binding.
KW SEQUENCE 1454 AA; 162716 MW; DEF8B2EFBD16F CRC64;

Query Match 61.8%; Score 34; DB 1; Length 1454;
Best Local Similarity 63.6%; Pred. No. 91;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 1307 RAINIGTNIH 1317

RESULT 13
TEGU HSV11 STANDARD; PRT; 3164 AA.
AC P10270;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Large tegument protein (Virion protein UL36).
GN UL36.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT Herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
CC -!- FUNCTION: Tegument protein.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EBV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X14112; CAA32311.1; -.
CC PIR; I30085; WMBEH6.
CC InterPro; IPR006928; Herpes teg N.
CC InterPro; IPR005210; Herpes_UL36.
CC Pfam; PF04843; Herpes teg N.
CC Pfam; PF03586; Herpes_UL36; 1.
KW Repeat.
FT DOMAIN 2911 2980 35 X 2 AA TANDEM REPEATS OF P-Q.
FT SEQUENCE 3164 AA; 335857 MW; CC5D31FF4F9FE3F4 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 3164;
Best Local Similarity 66.7%; Pred. No. 26+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQS:GTN 9
DB 1307 RAINIGTNIH 1317

```

```

Db      1401 RAEQALGTN 1409
RESULT 14
ID_YJHC_LACLA
YJHC_LACLA STANDARD; PRT; 87 AA.
AC QSCG26
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0237 protein yjhc.
GN YJHC OR L0946.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wicker P., Mauger S., Jaillon O., Malarre K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT Lactis ssp. lactis Il1403."
RL Genome Res. 11:731-753(2001).
CC -1- SIMILARITY: Belongs to the UPF0237 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB006329; AAK05044.1; -
CC PIR; B86743; B86743.
CC HAMAP; MF_01054; -; 1.
CC InterPro; IPR002912; ACT.
CC Pfam; PF01842; ACT; 1.
CC Hypothetical protein; Complete proteome.
SQ SEQUENCE 87 AA; 9282 NW; A7ED01657C149DE CRC64;-
Query Match 60.0%; Score 33; DB 1; Length 87;
Best Local Similarity 45.5%; Pred. No. 8.4;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 RASQSIGTNIH 11
Db 64 RGEALGVNIH 74
RESULT 15
KV3H_HUMAN STANDARD; PRT; 129 AA.
AC P04207;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region CLL precursor (Rheumatoid factor).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86177570; PubMed=3083417;
RA Jirik F.R., Sorge J., Fong S., Heitzmann J.G., Curd J.G., Chen P.P.,
RA Goldfien R., Carson D.A.;
RT "Cloning and sequence determination of a human rheumatoid factor
RT light-chain gene."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2195-2199(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12740; AAA58992.1; -
CC HSSP; P80362; 1WTL.
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0003823; F:antigen binding; NAS.
CC GO; GO:0006955; P:immune response; NAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_V.
CC Pfam; PF00047; Ig; 1.
CC SMART; SMO0406; IGV; 1.
CC PROSITE; PS50835; IG LIKE; 1.
CC Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 20 IG KAPPA CHAIN V-III REGION CLL.
FT DOMAIN 21 129 FRAMEWORK-1.
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 54 FRAMEWORK-2.
FT DOMAIN 55 69 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 70 76 FRAMEWORK-3.
FT DOMAIN 77 108 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 109 118 JKI SEGMENT.
FT DOMAIN 119 129 BY SIMILARITY.
FT DISULFID 43 108
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 14275 MW; 5C13B411BE60CC14 CRC64;
Query Match 60.0%; Score 33; DB 1; Length 129;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 RASQSIGTNI 10
Db 44 RASQSVSNL 53
Search completed: October 6, 2004, 16:30:27
Job time : 8.75439 secs

```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:59 ; Search time 11.5789 Seconds
(without alignments)
91.382 Million cell updates/sec

Title: US-09-635-974A-8

Perfect score: 55

Sequence: 1 RASQSIGTNIH 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 50 | 90.9 | 87 | 2 PH1082 | Ig kappa chain V r |
| 2 | 50 | 90.9 | 115 | 1 KVM5L7 | Ig kappa chain pre |
| 3 | 50 | 90.9 | 128 | 2 PN0445 | Ig kappa chain pre |
| 4 | 48 | 87.3 | 108 | 2 C30502 | Ig kappa chain V r |
| 5 | 44 | 80.0 | 114 | 2 S00996 | Ig kappa chain pre |
| 6 | 42 | 76.4 | 102 | 2 S26346 | Ig kappa chain V r |
| 7 | 42 | 76.4 | 104 | 2 B43413 | Ig kappa chain V r |
| 8 | 42 | 76.4 | 107 | 2 B45722 | anti-glycoprotein |
| 9 | 42 | 76.4 | 107 | 2 A45722 | anti-glycoprotein |
| 10 | 42 | 76.4 | 138 | 2 A26471 | Ig kappa chain pre |
| 11 | 41 | 74.5 | 107 | 2 C45722 | anti-glycoprotein |
| 12 | 40 | 72.7 | 95 | 2 PH0867 | Ig kappa chain V r |
| 13 | 39 | 70.9 | 91 | 2 S37511 | Ig kappa chain V r |
| 14 | 39 | 70.9 | 215 | 2 JEO243 | Ig kappa chain NIg |
| 15 | 38 | 69.1 | 122 | 2 S40370 | Ig kappa chain - h |
| 16 | 38 | 69.1 | 129 | 2 S40317 | Ig kappa chain - h |
| 17 | 38 | 69.1 | 144 | 2 B30502 | Ig heavy chain V r |
| 18 | 37 | 67.3 | 115 | 2 A25924 | Ig kappa chain pre |
| 19 | 37 | 67.3 | 127 | 2 S04577 | Ig kappa chain pre |
| 20 | 36 | 65.5 | 440 | 2 AD3173 | conserved hypothet |
| 21 | 36 | 65.5 | 886 | 2 T32229 | hypothetical prote |
| 22 | 35 | 63.6 | 86 | 2 S16827 | Ig kappa chain V r |
| 23 | 35 | 63.6 | 88 | 2 A37262 | Ig kappa chain V r |
| 24 | 35 | 63.6 | 116 | 2 B26555 | Ig kappa chain V-I |
| 25 | 35 | 63.6 | 163 | 2 S43762 | mannose-binding le |
| 26 | 35 | 63.6 | 255 | 2 T52352 | hypothetical prote |
| 27 | 35 | 63.6 | 261 | 2 S40162 | cathepsin G (EC 3. |
| 28 | 35 | 63.6 | 432 | 2 A97832 | hypothetical prote |
| 29 | 35 | 63.6 | 982 | 2 T06576 | probable protein k |

ALIGNMENTS

RESULT 1

PH1082

Ig light chain V region (clone 165.54) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000

C:Accession: PH1082

R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A>Title: Both IgM and Igg anti-DNA antibodies are the products of clonally selective B

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH1082

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-87 <TIL>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin

F16-80/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 90.9%; Score 50; DB 2; Length 87;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11

DB 14 RASQSIGTSH 24

RESULT 2

KVM5L7

Ig kappa chain precursor V region (L7) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 22-Jun-1999

C:Accession: A01925

R:Pech, M.; Hocht, J.; Schnell, H.; Zachau, H.G.

Nature 291, 668-670, 1981

A>Title: Differences between germ-line and rearranged immunoglobulin V-kappa coding seq

A:Reference number: A93259; MUID:81220975; PMID:6264318

A:Accession: A01925

A:Molecule type: DNA

A:Residues: 1-115 <PEC>

A:Cross-references: GB:V01564; GB:J00574; NID:G51718; PIDN:CAA24884.1; PID:G758153

A>Note: the sequence was determined from the germline gene

A>Note: there appear to be two possible splice junctions at the 3' end of the intron; th

C:Genetics:

A:Introns: 17/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai

hain disulfide bonds. In some cases, such as IGA and IgM, the subunits associate into la

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer

F11-20/Domain: signal sequence #status predicted <SIG>

F121-115/Product: Ig kappa chain V region (L7) #status predicted <MAT>

PH0867
IG kappa chain V region (anti-DNA, R3.5H5G) - human (fragment)
C:Species: Homo sapiens (man)
C:/Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 21-Jan-2000
C/Accession: PH0867
J. Manhemmel-Lory, A.; Katz, J.B.; Pillinger, M.; Chossein, C.; Smith, A.; Diamond, B.
R. Exp. Med. 174, 1639-1652, 1991
A:/Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiotypic
A:/Reference number: PH0862; MUID:92078875; PMID:1560528
A/Accession: PH0867
A:/Molecule type: DNA
A:/Residues: 1-95 <MAN>
C:/Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that bears
C:/Superfamily: immunoglobulin V region; immunoglobulin homology
C:/Keywords: heterotrimer; immunoglobulin
F:/1-23/Region: framework 1
F:/16-90/Domain: immunoglobulin homology <IMM>
F:/24-34/Region: complementarity-determining 1
F:/35-49/Region: framework 2
F:/50-56/Region: complementarity-determining 2
F:/57-88/Region: framework 3
F:/89-95/Region: complementarity-determining 3

```

Query Match      72.7%; Score 40; DB 2; Length 95;
Best Local Similarity 72.7%; Pred. No. 0.58;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGFLH 34

RESULT 13
S37511
Ig kappa chain V region (V-kappa 3) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37511
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IgM(+)IgG(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37511
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-91 <KLE>
A:Cross-references: EMBL:Z26600; NID:g405664; PIDN:CAA81354.1; PID:g405665
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match      70.9%; Score 39; DB 2; Length 91;
Best Local Similarity 70.0%; Pred. No. 0.89;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
DB 7 RASQSVGNL 16

RESULT 14
JE0243
Ig kappa chain NIG93 precursor - human
C:Species: Homo sapiens (man)
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0243
R:Alim, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazi, H.; T.
submitted to JIPID, November 1998
A:Description: A new subgroup of kappa type light chains (WkV) identified in cases of AL amy
A:Reference number: JE0243
A:Accession: JE0243
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match      70.9%; Score 39; DB 2; Length 215;
Best Local Similarity 70.0%; Pred. No. 2.3;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RASQSIGTNI 10
DB 24 RASQSVATNV 33

RESULT 15
S40370
Ig kappa chain - human
C:Species: Homo sapiens (man)
C>Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S40370
R:Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A>Title: Expressed human immunoglobulin chi genes and their hypermutation.
A:Reference number: S40312; MUID:94080891; PMID:8258341
A:Accession: S40370

```

```

A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-122 <KLE>
A:Cross-references: EMBL:X72480; NID:g441428; PIDN:CAA51148.1; PID:g441429
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:30-104/Domain: immunoglobulin homology <IMM>

Query Match      69.1%; Score 38; DB 2; Length 122;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 38 RASQSIGTFLH 48

Search completed: October 6, 2004, 16:34:51
Job time : 12.5789 secs

```

Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 9, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 113, App
Sequence 113, App
Sequence 113, App
Sequence 99, Appl
Sequence 113, App
Sequence 20, Appl
Sequence 112, App
Sequence 2, Appli
Sequence 19, Appl
Sequence 21, Appl

28 48 87.3 107 2 US-08-082-842A-1
29 48 87.3 109 1 US-07-942-245-4
30 47 85.5 11 2 US-08-232-081B-4
31 47 85.5 107 2 US-08-232-081B-9
32 47 85.5 107 2 US-08-232-081B-40
33 47 85.5 239 2 US-07-956-399-4
34 47 85.5 244 4 US-09-244-369B-1
35 45 81.8 106 1 US-08-326-362-4
36 44 80.0 96 1 US-08-276-852-113
37 44 80.0 96 1 US-08-898-575-113
38 44 80.0 96 1 US-08-899-575-113
39 44 80.0 96 4 US-09-472-087-99
40 44 80.0 96 5 PCT-US95-08743-113
41 44 80.0 155 4 US-09-472-087-20
42 44 80.0 155 4 US-09-472-087-112
43 43 78.2 145 3 US-09-096-244-2
44 42 76.4 100 1 US-08-436-463-19
45 42 76.4 103 1 US-08-436-463-21

ALIGNMENTS

RESULT 1
US-08-476-176B-53
; Sequence 53, Application US/08476176B
; Patent No. 5958708
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 5958708man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958708artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.176B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5958708ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-176B-53

Query Match 100.0%; Score 55; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 19,4912 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-8

Perfect score: 55

Sequence: 1 RASQSIGTNIH 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*

2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*

3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*

4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*

5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pcp.*

6: /cgn2_6/prodata/2/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 55 | 100.0 | 11 | 2 | US-08-476-176B-53 |
| 2 | 55 | 100.0 | 11 | 3 | US-08-127-721A-53 |
| 3 | 55 | 100.0 | 11 | 3 | US-08-485-246A-53 |
| 4 | 55 | 100.0 | 106 | 2 | US-08-800-198-4 |
| 5 | 55 | 100.0 | 106 | 3 | US-09-296-595-4 |
| 6 | 55 | 100.0 | 107 | 2 | US-08-476-176B-4 |
| 7 | 55 | 100.0 | 107 | 3 | US-08-127-721A-4 |
| 8 | 55 | 100.0 | 107 | 3 | US-08-485-246A-4 |
| 9 | 55 | 100.0 | 127 | 2 | US-08-476-176B-6 |
| 10 | 55 | 100.0 | 127 | 2 | US-08-476-176B-8 |
| 11 | 55 | 100.0 | 127 | 2 | US-08-476-176B-10 |
| 12 | 55 | 100.0 | 127 | 3 | US-08-127-721A-6 |
| 13 | 55 | 100.0 | 127 | 3 | US-08-127-721A-8 |
| 14 | 55 | 100.0 | 127 | 3 | US-08-127-721A-10 |
| 15 | 55 | 100.0 | 127 | 3 | US-08-485-246A-6 |
| 16 | 55 | 100.0 | 127 | 3 | US-08-485-246A-8 |
| 17 | 55 | 100.0 | 127 | 3 | US-08-485-246A-10 |
| 18 | 55 | 100.0 | 240 | 2 | US-08-800-198-8 |
| 19 | 55 | 100.0 | 240 | 3 | US-09-296-595-8 |
| 20 | 50 | 90.9 | 143 | 2 | US-08-653-402B-8 |
| 21 | 50 | 90.9 | 240 | 2 | US-07-956-399-2 |
| 22 | 48 | 87.3 | 11 | 1 | US-07-942-245-497 |
| 23 | 48 | 87.3 | 31 | 4 | US-08-525-539A-3 |
| 24 | 48 | 87.3 | 107 | 1 | US-08-436-463-20 |
| 25 | 48 | 87.3 | 107 | 1 | US-08-107-669B-1 |
| 26 | 48 | 87.3 | 107 | 1 | US-08-472-788A-1 |
| 27 | 48 | 87.3 | 107 | 2 | US-08-477-531B-1 |

Qy 1 RASQSIGTNIH 11
Db 1 RASQSIGTNIH 11

RESULT 2
US-08-127-721A-53
; Sequence 53, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066718artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,721A
; FILING DATE: 27-SEPTEMBER-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6066718ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-127-721A-53

Query Match 100.0%; Score 55; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RASQSIGTNIH 11
Db 1 RASQSIGTNIH 11

RESULT 3
US-08-485-246A-53
; Sequence 53, Application US/08485246A
; Patent No. 6072035
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6072035man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6072035artis Patent Department

STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-246A-53

Query Match 100.0%; Score 55; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RASQSIGTNIH 11
Db 1 RASQSIGTNIH 11

RESULT 4
US-08-800-198-4
; Sequence 4, Application US/08800198
; Patent No. 5942602
; GENERAL INFORMATION:
; APPLICANT: WELS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VARAKOLOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,198
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HAWLET-KING, DIANA
; REGISTRATION NUMBER: 33,302

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-127-721A-4

Query Match 100.0%; Score 55; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

RESULT 8
US-08-485-246A-4
Sequence 4, Application US/08485246A
Patent No. 6072035
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6072035man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6072035artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/485.246A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6072035ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-246A-4

Query Match 100.0%; Score 55; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

Db 24 RASQSIGTNIH 34

RESULT 9
US-08-476-176B-6
Sequence 6, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-176B-6

Query Match 100.0%; Score 55; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 44 RASQSIGTNIH 54

RESULT 10
US-08-476-176B-8
Sequence 8, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10

```

/ FILING DATE: 25-SEPTEMBER-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5958708ak, Henry P.
/ REGISTRATION NUMBER: 33,200
/ REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 277-5110
/ TELEFAX: (908) 277-4306
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 127 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-476-1765-10
/
Query Match 100.0%; Score 55; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 12
US-08-127-721A-6
/ Sequence 6, Application US/08127721A
/ Patent No. 6066718
/ GENERAL INFORMATION:
/ APPLICANT: Hardman, No. 6066718man
/ APPLICANT: Kolbinger, Frank
/ APPLICANT: Saldanha, Jose
/ TITLE OF INVENTION: Reshaped monoclonal antibodies against an
/ TITLE OF INVENTION: Reshaped monoclonal antibodies against an
/ NUMBER OF SEQUENCES: 55
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: No. 6066718artis Patent and Trademark Department
/ STREET: 59 Route 10
/ CITY: East Hanover
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07936-1080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/127,721A
/ FILING DATE: 27-SEPTEMBER-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/952,802
/ FILING DATE: 25-SEPTEMBER-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 6066718ak, Henry P.
/ REGISTRATION NUMBER: 33,200
/ REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (908) 277-5110
/ TELEFAX: (908) 277-4306
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 127 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-127-721A-6
/
Query Match 100.0%; Score 55; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 13
US-08-127-721A-8
; Sequence 8, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6066718artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 25-SEPTEMBER-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6066718ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-127-721A-8

Query Match 100.0%; Score 55; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 14
US-08-127-721A-10
; Sequence 10, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6066718artis Patent and Trademark Department

STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6066718ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-127-721A-10

Query Match 100.0%; Score 55; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

RESULT 15
US-08-485-246A-6
; Sequence 6, Application US/08485246A
; Patent No. 6072035
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6072035man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6072035artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992

```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6072035ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-246A-6

Query Match 100.0%; Score 55; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

```

Search completed: October 6, 2004, 16:36:33
Job time : 20.4912 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 55.5789 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-8
Perfect score: 55
Sequence: 1 RASQSIGTNIH 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan94:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|---------------------|
| 1 | 55 | 100.0 | 11 | 2 AAY26990 | CDRI doma |
| 2 | 55 | 100.0 | 11 | 3 AAY59314 | Aay59314 Light cha |
| 3 | 55 | 100.0 | 11 | 4 AAB37955 | Aab37955 Anti-EGFR |
| 4 | 55 | 100.0 | 11 | 5 AAU77787 | Aau77787 Mouse lig |
| 5 | 55 | 100.0 | 106 | 2 AAW71241 | Aaw71241 Light cha |
| 6 | 55 | 100.0 | 107 | 2 AAR50190 | Aar50190 Light cha |
| 7 | 55 | 100.0 | 107 | 2 AAW08949 | Aaw08949 Kappa lig |
| 8 | 55 | 100.0 | 107 | 2 AAW08948 | Aaw08948 Kappa lig |
| 9 | 55 | 100.0 | 107 | 2 AAY26979 | Aay26979 Light cha |
| 10 | 55 | 100.0 | 107 | 3 AAY70604 | Aay70604 V kappa re |
| 11 | 55 | 100.0 | 127 | 2 AAR50192 | Aar50192 Light cha |
| 12 | 55 | 100.0 | 127 | 2 AAR50187 | Aar50187 Light cha |
| 13 | 55 | 100.0 | 127 | 2 AAR50191 | Aar50191 Light cha |
| 14 | 55 | 100.0 | 127 | 2 AAW08945 | Aaw08945 Kappa lig |
| 15 | 55 | 100.0 | 127 | 2 AAW08941 | Aaw08941 Kappa lig |
| 16 | 55 | 100.0 | 127 | 2 AAW08946 | Aaw08946 Kappa lig |
| 17 | 55 | 100.0 | 127 | 2 AAW08943 | Aaw08943 Kappa lig |
| 18 | 55 | 100.0 | 127 | 2 AAY26981 | Aay26981 Light cha |
| 19 | 55 | 100.0 | 127 | 2 AAY26980 | Aay26980 Light cha |
| 20 | 55 | 100.0 | 127 | 2 AAY26982 | Aay26982 Light cha |
| 21 | 55 | 100.0 | 240 | 2 AAW05133 | Aaw05133 Single ch |
| 22 | 55 | 100.0 | 240 | 2 AAW71243 | Aaw71243 scfv comp |
| 23 | 55 | 100.0 | 245 | 3 AAY70605 | Aay70605 scfv frag |
| 24 | 55 | 100.0 | 651 | 3 AAW05135 | Aaw05135 scfv (225) |
| 25 | 55 | 100.0 | 892 | 2 AAW05140 | Aaw05140 scfv2(225) |

| | | | | | | |
|----|----|-------|------|---|-----------|------------|
| 26 | 55 | 100.0 | 892 | 2 | AAW05139 | scFv2(FRP |
| 27 | 55 | 100.0 | 1020 | 2 | AAW05141 | scFv2(FRP |
| 28 | 50 | 90.9 | 11 | 2 | AAW44178 | Monoclonal |
| 29 | 50 | 90.9 | 11 | 5 | AAU72851 | Anti-NGK2 |
| 30 | 50 | 90.9 | 107 | 2 | AAAR32129 | Anti-IL2R |
| 31 | 50 | 90.9 | 107 | 5 | AAU72850 | Anti-NGK2 |
| 32 | 50 | 90.9 | 127 | 2 | AAW44176 | Monoclonal |
| 33 | 50 | 90.9 | 143 | 2 | AAW19580 | Mouse ant |
| 34 | 50 | 90.9 | 214 | 6 | ABP96757 | TSH recep |
| 35 | 50 | 90.9 | 214 | 6 | ABP96766 | TSH recep |
| 36 | 50 | 90.9 | 214 | 6 | ABP96758 | TSH recep |
| 37 | 50 | 90.9 | 214 | 6 | ABP96762 | TSH recep |
| 38 | 50 | 90.9 | 214 | 6 | ABP96761 | TSH recep |
| 39 | 50 | 90.9 | 214 | 6 | ABP96765 | TSH recep |
| 40 | 50 | 90.9 | 510 | 5 | AAU72860 | Human p53 |
| 41 | 49 | 89.1 | 11 | 3 | AAAB10013 | H. pylori |
| 42 | 49 | 89.1 | 11 | 4 | AAAB8093 | H. pylori |
| 43 | 49 | 89.1 | 11 | 4 | AAAB8061 | H. pylori |
| 44 | 49 | 89.1 | 107 | 3 | AAAB10023 | H. pylori |
| 45 | 49 | 89.1 | 107 | 4 | AAAB86109 | H. pylori |

ALIGNMENTS

RESULT 1
AAY26990
ID AAY26990 standard; protein; 11 AA.
XX AC AAY26990;
XX AC AAY26990;

DT 24-DEC-1999 (first entry)

DE CDRI domain reshaped human antibody light chain variable region.

XX Diagnosis: IGE; immunoglobulin; body fluid; human; monoclonal antibody;
KW binding affinity; mouse; CDR; complementarity determining region;
KW allergy.

OS Synthetic.

OS Homo sapiens.

XX US9595708-A.

XX 28-SEP-1999.

XX 07-JUN-1995; 95US-00476176.

XX 25-SEP-1992; 92US-00952802.

XX 27-SEP-1993; 93US-00127721.

XX (NOVS) NOVARTIS CORP.

XX (TANO-) TANOX BIOSYSTEMS INC.

XX Saldanha J, Kolbinger F, Hardman N;

XX WPI; 1999-570765/48.

XX new method for determining Igs levels in a sample.

XX Claim 5; Col 67; 19pp; English.

XX The invention relates to a method of determining IGE levels in a body
fluid sample, by contacting the sample with a reshaped human monoclonal
antibody (RA) having a binding affinity about equal to that of the murine
CDR-donor antibody TES-C21 produced by the cell line 11133. The
antibodies are useful in the diagnosis, prophylaxis and treatment of
allergy. This sequence represents the complementarity determining region
(CDR) 1 from the light chain variable domain of a reshaped human antibody

SQ Sequence 11 AA;

Query Match 100.0%; Score 55; DB 2; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 1 RASQSIGTNIH 11

RESULT 2
AA59314
ID AA59314 standard; peptide; 11 AA.
AC AA59314;
DT 07-MAR-2000 (first entry)
XX Light chain hypervariable region, CDR1.
DE Hypervariable region; complementarity determining region; CDR; tumour;
KW single chain antibody; growth inhibitor; human; tumourigenesis; therapy;
KW protein receptor tyrosine kinase; light chain.
XX
OS Mus sp.
XX
XX WO9960023-A1.
XX
XX 25-NOV-1999.
XX
XX 14-MAY-1999; 99WO-US010741.
XX
XX 15-MAY-1998; 98US-00079612.
XX 15-MAY-1998; 98US-0085613P.
XX 07-DEC-1998; 98US-00206138.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (UASR-) UAB RES FOUND.
XX
XX Waksal HW, Saleh MN, Robert F, Buchsbaum DJ;
XX
XX WPI; 2000-062440/05.
XX N-PSDB; AA248628.
XX
XX Treatment of human tumors, using a combination of radiation and a non-
XX radiolabeled protein receptor tyrosine kinase inhibitor.
XX
XX Disclosure; Page 15; 31pp; English.
XX
XX This sequence is the hypervariable region CDR1 (complementarity
XX determining region 1) of the light chain of a single chain antibody
XX derived from the murine antibody 225. The invention relates to a method
XX for inhibiting the growth of tumors in human patients by treating with
XX an effective amount of a combination of radiation and a non-radiolabelled
XX protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
XX which can lead to tumourigenesis. The method can be used in the treatment
XX of tumors of e.g. breast, lung, colon, kidney, bladder, head and neck,
XX ovary, prostate or brain. The administration of a suitable antibody to
XX the patient makes the tumour more susceptible to radiotherapy
XX
XX Sequence 11 AA;
XX
XX Query Match 100.0%; Score 55; DB 3; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.00023;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 1 RASQSIGTNIH 11

RESULT 3
AAB37955
ID AAB37955 standard; protein; 11 AA.
XX

AC AAB37955;
XX
XX 12-MAR-2001 (first entry)
XX
XX Anti-EGFR monoclonal antibody L chain V region CDR1 peptide sequence.
XX
XX Refractory tumour growth inhibition; epidermal growth factor receptor;
KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
KW complementarity determining region; CDR.
XX
XX Mus sp.
XX
XX WO200069459-A1.
XX
XX 23-NOV-2000.
XX
XX 01-MAY-2000; 2000WO-US011756.
XX
XX 14-MAY-1999; 99US-00312284.
XX 13-AUG-1999; 99US-00374028.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX
XX Waksal HW;
XX
XX WPI; 2001-016160/02.
XX N-PSDB; AAC83238.
XX
XX Epidermal growth factor receptor/human epidermal growth factor receptor-1
XX antagonist for inhibiting the growth of refractory tumors.
XX
XX Disclosure; Page 14; 31pp; English.
XX
XX This invention relates to a method for inhibiting the growth of
XX refractory tumors that are stimulated by a ligand of epidermal growth
XX factor receptor (EGFR) in human patients. The method involves treating
XX the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
XX optionally with a chemotherapeutic agent or radiation. The antagonist can
XX be for example a chimeric anti-EGFR monoclonal antibody, C225. The
XX EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
XX tumors such as tumors of breast, heart, lung, small intestine, colon,
XX spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas,
XX skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
XX liver, preferably squamous cell carcinomas. The present sequence
XX represents the light chain variable region complementarity determining
XX region 1 amino acid sequence of the chimeric anti-EGFR monoclonal
XX antibody C225 which is used in an example illustrating the method of the
XX invention
XX
XX Sequence 11 AA;
XX
XX Query Match 100.0%; Score 55; DB 4; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.00023;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 1 RASQSIGTNIH 11

RESULT 4
AAU77787
ID AAU77787 standard; peptide; 11 AA.
XX
XX AAU77787;
XX
XX 05-JUN-2002 (first entry)
XX
XX Mouse light chain hypervariable region (CDR1) of 225 antibody.
XX
XX Mouse; light chain; antibody; hyperproliferative disease;
KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
KW seborrheic keratosis; warts; keloid scars; eczema; 255 antibody;

KW hypervariable region; CDRI; EGFR inhibitor.
 XX Mus sp.
 OS
 XX WO200211677-A2.
 PN
 XX 14-FEB-2002.
 PD
 XX 09-AUG-2001; 2001WO-US041647.
 PF
 XX 09-AUG-2000; 2000US-00635974.
 PR
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA
 XX Teufel T;
 PI
 XX WPI; 2002-257423/30.
 DR
 XX N-PSDB; ABK11443.
 DR
 XX Treating a mammal with hyperproliferative disease especially psoriasis,
 PT stimulated by ligand of member of epidermal growth factor family of
 PT receptors, by administering antagonist of the receptor.
 PT
 XX Disclosure; Page 12; 28pp; English.
 PS
 XX This invention relates to a novel method for treating a mammal with
 CC hyperproliferative disease stimulated by a ligand of a member of the
 CC epidermal growth factor (EGF) family of receptors. The method involves
 CC administering an antibody or a defective receptor that is an antagonist
 CC of a member of the EGF receptor family, or a combination of the
 CC antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
 CC The antibody used in the method of the invention acts as an epidermal
 CC growth factor receptor (EGFR) antagonist by inhibiting EGFR/HER1
 CC phosphorylation. The method of the invention is useful for treating a
 CC mammal with hyperproliferative disease such as psoriasis, actinic
 CC keratoses, seboreic keratoses, warts, keloid scars and eczema
 CC stimulated by a ligand of a member of the EGF family of receptor. This
 CC sequence represents the murine anti-EGF 255 antibody light chain (LH)
 CC hypervariable region (CDRI) used as an inhibitor of EGFR in the method of
 CC the invention
 CC
 XX
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 55; DB 5; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RASQSIGTNIH 11
 |||||
 DB 1 RASQSIGTNIH 11
 |||||
 RESULT 5
 AAW71241
 ID AAW71241 standard; protein; 106 AA.
 XX
 AC AAW71241;
 XX
 XX 20-NOV-1998 (first entry)
 DT
 XX
 DE Light chain variable domain of antibody 14E1.
 XX
 KW Light chain variable domain; antibody 14E1;
 KW human epidermal growth factor receptor; EGF-R; treatment; carcinoma;
 KW EGF-Rviii; in vitro purging; breast carcinoma; ss.
 XX
 XX Mus sp.
 OS
 XX WO9836074-A2.
 PN
 XX 20-AUG-1998.
 PD
 XX 13-FEB-1998; 98WO-IB000413.
 PF

XX 13-FEB-1997; 97US-00800198.
 PR
 XX (SCHD) SCHERING AG.
 PA
 XX Weis WS, Schmidt M, Vakalopoulou E, Schneider DW;
 PI
 XX WPI; 1998-467176/40.
 DR
 XX N-PSDB; AAV54788.
 DR
 XX New single chain polypeptide that binds to epidermal growth factor
 PT receptor - is derived from antibody 14E, used for treatment and diagnosis
 PT of cancer or other cell proliferative diseases.
 PT
 XX Example 1; Fig 11D; 48pp; English.
 PS
 XX The present sequence represents the light chain variable domain of murine
 CC monoclonal antibody 14E1. This antibody is active against the human
 CC epidermal growth factor receptor (EGF-R). The sequence is used to
 CC construct a single-chain polypeptide (scFv) that has binding affinity for
 CC an EGF-R. The scFv comprises two segments which contain the binding parts
 CC of the heavy and light chain variable domains of the monoclonal antibody
 CC 14E1. The scFvs of the invention are used to treat carcinomas (gliomas,
 CC melanomas and tumours) that overexpress EGF-R, specifically EGF-Rviii,
 CC relative to normal cells, particularly by blocking proliferation of such
 CC cells by preventing receptor activation. They can also be used for in
 CC vitro purging. When the scFv is appropriately labelled, it can be used
 CC diagnostically to image cells that express EGF-R, specifically breast
 CC carcinoma
 CC
 XX
 SQ Sequence 106 AA;
 Query Match 100.0%; Score 55; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RASQSIGTNIH 11
 |||||
 DB 24 RASQSIGTNIH 34
 |||||
 RESULT 6
 AAR50190
 ID AAR50190 standard; protein; 107 AA.
 XX
 AC AAR50190;
 XX
 XX 25-MAR-2003 (revised)
 DT 18-OCT-1994 (first entry)
 DE
 XX Light chain variable region of TES-C21 antibody.
 XX
 KW Antibody; heavy chain; light chain; variable region; therapy;
 KW constant region; prophylaxis; prevention; allergy; identification;
 KW allergic reaction; immunoglobulin; IgE; determination.
 XX
 XX Mus musculus.
 OS
 XX Key Location/Qualifiers
 FH Binding-site 24..34
 FT Binding-site /label= CDR 1.
 FT Binding-site 50..56
 FT Binding-site /label= CDR 2.
 FT Binding-site 89..97
 FT Binding-site /label= CDR 3.
 XX
 XX EP589840-A1.
 PN
 XX 30-MAR-1994.
 PD
 XX 15-SEP-1993; 93EP-00810653.
 PF
 XX 24-SEP-1992; 92GB-00020228.
 PR

```

PR 25-SEP-1992; 92US-00952802.
XX (CIBA ) CIBA GEIGY AG.
PA (TANO-) TANOX BIOSYSTEMS INC.
PA (NOVS ) NOVARTIS AG.
XX
XX Hardman N, Kolbinger F, Saldanha J;
XX WPI; 1994-103410/13.
XX N-PSDB; AAQ44714.
XX
XX New reshaped human monoclonal antibody specific for IgE - used for
PT prophylaxis or treatment of allergic reactions or qualitative or
PT quantitative determ. of IgE.
XX
XX Disclosure; Page 28; 68pp; English.
XX
XX Reshaped human monoclonal antibodies can be used in the prophylaxis
CC and/or treatment of allergic reactions in humans. The monoclonal
CC antibodies can also be used for the qualitative or quantitative
CC determination of IgE and for the determination of surface IgE positive B
CC cells. They can provide a long lasting therapeutic effect without
CC inducing immunogenicity as foreign proteins. The CDR's from the variable
CC heavy and light chains of the murine antibody TES-C21 were used in the
CC construction of a molecular model of the variable heavy and light chains
CC of a human reshaped antibody which also comprised human variable heavy
CC and light chain framework regions. (Updated on 25-MAR-2003 to correct PN
CC field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 107 AA;
SQ
Query Match 100.0%; Score 55; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

RESULT 7
AAW08949
ID AAW08949 standard; protein; 107 AA.
XX
XX AAW08949;
XX
XX 18-SEP-1997 (first entry)
XX
XX Kappa light chain variable region of 225RB antibody.
XX
XX Kappa; light chain; reshaped; monoclonal; antibody; 225RB; human;
XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX late stage; prostate; prostate; variable region; framework;
XX complementarity determining region; CDR.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..23 /label= framework_1
FT Region 24..34 /label= CDR_1
FT Region 35..49 /label= framework_2
FT Region 50..56 /label= CDR_2
FT Region 57..88 /label= framework_3
FT Region 89..97 /label= CDR_3
FT Region 98..107 /label= framework_4
XX

PN WO9640210-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009847.
XX
XX 07-JUN-1995; 95US-00482982.
XX 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX WPI; 1997-051897/05.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Claim 29; Fig 21; 112pp; English.
XX
XX The present sequence is the kappa light chain variable region of the
CC reshaped human monoclonal antibody (Mab) H225, 225SKA. The Mab is
CC specific for the human epidermal growth factor (EGF) receptor. The Mab,
CC or a fragment, can be used to inhibit the growth of tumour cells,
CC especially late stage prostatic tumour cells in humans, optionally
CC conjugated to a cytotoxic agent, especially doxorubicin, taxol or
CC cisplatin, or a signal transduction, ras or cell cycle inhibitor
XX
XX Sequence 107 AA;
SQ
Query Match 100.0%; Score 55; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34

RESULT 8
AAW08948
ID AAW08948 standard; protein; 107 AA.
XX
XX AAW08948;
XX
XX 18-SEP-1997 (first entry)
XX
XX Kappa light chain variable region of 225RA antibody.
XX
XX Kappa; light chain; reshaped; monoclonal; antibody; 225RA; human;
XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX late stage; prostate; prostate; variable region; framework;
XX complementarity determining region; CDR.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..23 /label= framework_1
FT Region 24..34 /label= CDR_1
FT Region 35..49 /label= framework_2
FT Region 50..56 /label= CDR_2
FT Region 57..88 /label= framework_3
FT Region 89..97 /label= CDR_3
FT Region 98..107 /label= framework_4
XX

```

```

PN WO9640210-A1.
XX
XX
PD 19-DEC-1996.
XX
XX
PF 07-JUN-1996; 96WO-US0009847.
XX
XX
PR 07-JUN-1995; 95US-00482982.
XX
XX
PR 15-DEC-1995; 95US-00573289.
XX
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
PI
XX WPI; 1997-051897/05.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Claim 29; Fig 21; 112pp; English.
XX
XX The present sequence is the kappa light chain variable region of the
CC reshaped human monoclonal antibody (MAB) H225, 225RKA. The MAB is
CC specific for the human epidermal growth factor (EGF) receptor. The MAB,
CC or a fragment, can be used to inhibit the growth of tumour cells,
CC especially late stage prostatic tumour cells in humans, optionally
CC conjugated to a cytotoxic agent, especially doxorubicin, taxol or
CC cisplatin, or a signal transduction, ras or cell cycle inhibitor
XX
XX Sequence 107 AA;
SQ
Query Match 100.0%; Score 55; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
DB 24 RASQSIGTNIH 34
|||||
RESULT 10
AAAY70604
ID AAY70604 standard; protein; 107 AA.
XX
XX AC AAY70604;
XX
XX 18-JUL-2000 (first entry)
DT
XX
DE Kappa region of humanised antibody Hu-901.
XX
KW Humanised antibody-901; Hu-901; V kappa; kappa chain variable region;
KW anti-IgE; anti-immunoglobulin E; monoclonal antibody; hybridoma TES-C21;
KW recombinant adenovirus vector; anti-IgE antibody; gene therapy; scFv;
KW single chain variable fragment; allergy; antiallergic.
XX
XX Homo sapiens.
OS
XX Mus sp.
XX
XX WO200015260-A1.
XX
XX 23-MAR-2000.
XX
XX 16-SEP-1999; 99WO-US021646.
PF
XX 16-SEP-1998; 98US-0100639P.
PR
XX (TANO-) TANOX INC.
PA
XX Liou RS, Thomas D;
PI
XX WPI; 2000-271264/23.
XX
XX N-PSDB; AAZ52078.
XX
XX New nucleic acids encoding anti-IgE antibodies such as Hu-901 or their
PT fragments, useful in the gene therapy of allergic diseases.
XX
XX Disclosure; Page 47; 52pp; English.
XX
XX The present protein sequence is the V kappa (kappa chain variable region)
CC of the humanised antibody Hu-901. This is an exemplary anti-IgE (anti-
CC immunoglobulin E) monoclonal antibody produced by hybridoma TES-C21. The
CC heavy and light chains are obtained through RT-PCR using the RNA from the
CC transfectoma cell line producing the chimeric antibody. Humanised anti-
CC IgE genes are incorporated in the genome of a recombinant adenovirus
CC vector as an independent transcriptional unit, and packaged into
CC infectious virus particles. Upon infection of host, the recombinant
CC adenovirus vector will direct the production of either intact anti-IgE
CC antibody or an scFv fragment in the serum, which will bind free

```

CC circulating IgE, resulting in the reduction of free serum IgE. The
 CC binding of the antibody or fragment to IgE-bearing B cells may lower IgE
 CC levels by down-regulating IgE production by these B cells. These methods
 CC are useful in the gene therapy of allergic diseases
 XX
 SQ Sequence 107 AA;

Query Match 100.0%; Score 55; DB 3; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 24 RASQSIGTNIH 34

RESULT 11
 AAR50192
 ID AAR50192 standard; protein; 127 AA.

XX AAR50192;
 XX
 DT 25-MAR-2003 (revised)
 DT 18-OCT-1994 (first entry)
 XX
 DE Light chain variable region C21-L3 of reshaped antibody.
 XX
 KW Antibody; heavy chain; light chain; variable region; therapy;
 KW constant region; prophylaxis; prevention; allergy; identification;
 KW allergic reaction; immunoglobulin; IgE; determination.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Binding-site 44..54
 FT /label= CDR 1.
 FT Binding-site 70..76
 FT /label= CDR 2.
 FT Binding-site 109..117
 FT /label= CDR 3.

XX
 PN EF589840-A1.
 XX
 PD 30-MAR-1994.
 XX
 PF 15-SEP-1993; 93EP-00810653.
 XX
 PR 24-SEP-1992; 92GB-00020228.
 PR 25-SEP-1992; 92US-00952802.
 XX
 PA (CIBA) CIBA GEIGY AG.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PA (NOVS) NOVARTIS AG.
 XX
 PI Hardman N, Kolbinger F, Saldanha J;

XX
 DR WPI; 1994-103410/13.
 DR N-PSDB; AAQ44716.
 XX
 PT New reshaped human monoclonal antibody specific for IgE - used for
 PT prophylaxis or treatment of allergic reactions or qualitative or
 PT quantitative determin. of IgE.
 XX
 PS Example 2; Page 36; 68pp; English.

XX Reshaped human monoclonal antibodies can be used in the prophylaxis
 CC and/or treatment of allergic reactions in humans. The monoclonal
 CC antibodies can also be used for the qualitative or quantitative
 CC determination of IgE and for the determination of surface IgE positive B
 CC cells. They can provide a long lasting therapeutic effect without
 CC inducing immunogenicity as foreign proteins. This sequence is a slightly
 CC modified version of C21-L1 described in AAR50187, having glutamic acid at
 CC position 1 of the mature protein instead of aspartic acid and valine at

CC position 3 of the mature protein instead of leucine (positions 21 and 23
 CC respectively of this sequence). (Updated on 25-MAR-2003 to correct PN
 CC field.) (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 127 AA;

Query Match 100.0%; Score 55; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 44 RASQSIGTNIH 54

RESULT 12
 AAR50187
 ID AAR50187 standard; protein; 127 AA.

XX AAR50187;
 XX
 DT 25-MAR-2003 (revised)
 DT 18-OCT-1994 (first entry)
 XX
 DE Light chain variable region C21-L1 of reshaped antibody.
 XX
 KW Antibody; heavy chain; light chain; variable region; therapy;
 KW constant region; prophylaxis; prevention; allergy; identification;
 KW allergic reaction; immunoglobulin; IgE; determination.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Binding-site 44..54
 FT /label= CDR 1.
 FT Binding-site 70..76
 FT /label= CDR 2.
 FT Binding-site 109..117
 FT /label= CDR 3.

XX
 PN EF589840-A1.
 XX
 PD 30-MAR-1994.
 XX
 PF 15-SEP-1993; 93EP-00810653.
 XX
 PR 24-SEP-1992; 92GB-00020228.
 PR 25-SEP-1992; 92US-00952802.
 XX
 PA (CIBA) CIBA GEIGY AG.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 PA (NOVS) NOVARTIS AG.
 XX
 PI Hardman N, Kolbinger F, Saldanha J;

XX
 DR WPI; 1994-103410/13.
 DR N-PSDB; AAQ44711.
 XX
 PT New reshaped human monoclonal antibody specific for IgE - used for
 PT prophylaxis or treatment of allergic reactions or qualitative or
 PT quantitative determin. of IgE.

XX Claim 4; Page 29-30; 68pp; English.

XX Reshaped human monoclonal antibodies can be used in the prophylaxis
 CC and/or treatment of allergic reactions in humans. The monoclonal
 CC antibodies can also be used for the qualitative or quantitative
 CC determination of IgE and for the determination of surface IgE positive B
 CC cells. They can provide a long lasting therapeutic effect without
 CC inducing immunogenicity as foreign proteins. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 127 AA;

Query Match 100.0%; Score 55; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 44 RASQSIGTNIH 54

RESULT 13
 ID AAR50191
 ID AAR50191 standard; protein; 127 AA.
 XX AC AAR50191;
 XX DT 25-MAR-2003 (revised)
 XX DT 18-OCT-1994 (first entry)
 XX DE Light chain variable region C21-L2 of reshaped antibody.
 XX XX Antibody; heavy chain; light chain; variable region; therapy;
 KW constant region; prophylaxis; prevention; allergy; identification;
 KW allergic reaction; immunoglobulin; IgE; determination.
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Binding-site 44..54
 FT /label= CDR 1.
 FT Binding-site 70..76
 FT /label= CDR 2.
 FT Binding-site 109..117
 FT /label= CDR 3.

XX PF EP589840-A1.

XX PN 30-MAR-1994.

XX PD 15-SEP-1993; 93EP-00810653.

XX PR 24-SEP-1992; 92GB-00020228.

XX PR 25-SEP-1992; 92US-00952802.

XX PA (CIBA) CIBA GEIGY AG.

XX PA (TANO-) TANOX BIOSYSTEMS INC.

XX PA (NOVS) NOVARTIS AG.

XX PI Hardman N, Kolbinger F, Saldanha J;

XX DR WPI; 1994-103410/13.

XX DR N-PSDB; AAQ44715.

XX PT New reshaped human monoclonal antibody specific for IgE - used for

PT prophylaxis or treatment of allergic reactions or qualitative or

PT quantitative determ. of IgE.

XX PS Example 2; Page 33; 68pp; English.

XX CC Reshaped human monoclonal antibodies can be used in the prophylaxis
 CC and/or treatment of allergic reactions in humans. The monoclonal
 CC antibodies can also be used for the qualitative or quantitative
 CC determination of IgE and for the determination of surface IgE positive B
 CC cells. They can provide a long lasting therapeutic effect without
 CC inducing immunogenicity as foreign proteins. This sequence is a slightly
 CC modified version of C21-L1 described in AAR50187, having aspartic acid at
 CC position 50 of the mature protein (position 80 of this sequence) instead
 CC of serine. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-
 CC MAR-2003 to correct PA field.)

XX SQ Sequence 127 AA;

Query Match 100.0%; Score 55; DB 2; Length 127;

Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 44 RASQSIGTNIH 54

RESULT 14
 ID AAW08945
 ID AAW08945 standard; protein; 127 AA.
 XX AC AAW08945;
 XX DT 18-SEP-1997 (first entry)
 XX DE Kappa light chain variable region of C225 antibody.

XX KW Kappa; light chain; murine; mouse; monoclonal; antibody; C225; human;
 KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 KW late stage; prostatic; prostate; variable region; chimeric; L7'CL;
 KW leader sequence.

XX OS Mus; spp.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.

XX PN WO9640210-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009847.

XX PR 07-JUN-1995; 95US-00482982.

XX PR 15-DEC-1995; 95US-00573289.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PA (MRCC-) MRC COLLABORATIVE CENT.

XX PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;

XX DR WPI; 1997-051897/05.

XX DR N-PSDB; AAT49344.

XX PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 PT for inhibiting tumour growth, esp. of late stage prostatic tumour.

XX PS Example IV; Fig 17; 112pp; English.

XX CC The present sequence is the kappa light chain variable region of the
 CC chimeric monoclonal antibody (Mab) C225, with the modified leader
 CC sequence from the kappa light chain of L7'CL Mab. C225 is specific for
 CC the human epidermal growth factor (EGF) receptor. C225, or a fragment,
 CC can be used to inhibit the growth of tumour cells, especially late stage
 CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 CC agent, especially doxorubicin, taxol or cisplatin, or a signal
 CC transduction, ras or cell cycle inhibitor

XX SQ Sequence 127 AA;

Query Match 100.0%; Score 55; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RASQSIGTNIH 11
 |||||
 Db 44 RASQSIGTNIH 54

RESULT 15
 ID AAW08941
 ID AAW08941 standard; protein; 127 AA.

```

XX
AC AAW08941;
XX
XX 27-AUG-2003 (revised)
DT 18-SEP-1997 (first entry)
XX
DE Kappa light chain variable region of M225 antibody.
XX
KW Kappa; light chain; murine; mouse; monoclonal; antibody; M225; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region.
XX
OS Mus sp.
XX
XX WO9640210-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009847.
XX
XX 07-JUN-1995; 95US-00482982.
XX 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
XX WPI; 1997-051897/05.
XX DR N-PSDB; AAT49338.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Example IV; Fig 13; 112pp; English.
XX
XX The present sequence is the kappa light chain variable region of the
CC murine monoclonal antibody (Mab) M225, which is specific for the human
CC epidermal growth factor (EGF) receptor. The Mab, or a fragment, can be
CC used to inhibit the growth of tumour cells, especially late stage
CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
CC agent, especially doxorubicin, taxol or cisplatin, or a signal
CC transduction, ras or cell cycle inhibitor. (Updated on 27-AUG-2003 to
CC correct OS field.)
XX
XX Sequence 127 AA;
XX
Query Match 100.0%; Score 55; DB 2; Length 127;
Best Local Similarity 100.0%; Pred.No. 0.0039;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RASQSIGTNIH 11
Db 44 RASQSIGTNIH 54

```

Search completed: October 6, 2004, 16:29:46
Job time : 56.5789 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 37.3684 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-10

Perfect score: 25

Sequence: 1 ASESIS 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 25 | 100.0 | 6 | 10 | US-09-798-689-34 |
| 2 | 25 | 100.0 | 6 | 10 | US-09-998-954B-10 |
| 3 | 25 | 100.0 | 7 | 12 | US-10-239-656-30 |
| 4 | 25 | 100.0 | 7 | 12 | US-10-239-656-40 |
| 5 | 25 | 100.0 | 7 | 12 | US-10-374-600-101 |
| 6 | 25 | 100.0 | 7 | 15 | US-10-374-531-101 |
| 7 | 25 | 100.0 | 19 | 14 | US-10-029-386-30831 |
| 8 | 25 | 100.0 | 46 | 12 | US-10-424-599-191046 |
| 9 | 25 | 100.0 | 61 | 16 | US-10-379-151-3 |
| 10 | 25 | 100.0 | 61 | 16 | US-10-379-151-18 |
| 11 | 25 | 100.0 | 88 | 12 | US-10-424-599-229773 |
| 12 | 25 | 100.0 | 107 | 9 | US-09-991-470-25 |
| 13 | 25 | 100.0 | 107 | 12 | US-10-239-656-28 |
| 14 | 25 | 100.0 | 107 | 12 | US-10-239-656-38 |
| 15 | 25 | 100.0 | 107 | 12 | US-10-453-696-129 |

Sequence 113, App
Sequence 114, App
Sequence 10, Appl
Sequence 129, App
Sequence 4, Appl
Sequence 113, App
Sequence 129, App
Sequence 20, Appl
Sequence 22, Appl
Sequence 23, Appl
Sequence 22, Appl
Sequence 23, Appl
Sequence 5, Appl
Sequence 11, Appl
Sequence 15, Appl
Sequence 17, Appl
Sequence 57, Appl
Sequence 11, Appl
Sequence 53273, A
Sequence 45496, A
Sequence 27, Appl
Sequence 5606, Ap
Sequence 1543, Ap
Sequence 8, Appl

US-10-374-600-113
US-10-374-600-114
US-10-091-236-10
US-10-308-817-129
US-10-379-189-4
US-10-374-531-113
US-10-374-531-114
US-10-412-703A-129
US-10-374-600-20
US-10-374-600-22
US-10-374-600-23
US-10-374-531-20
US-10-374-531-22
US-10-374-531-23
US-10-374-600-5
US-10-374-600-11
US-10-374-600-15
US-10-374-600-17
US-10-374-531-5
US-10-374-531-11
US-10-374-531-15
US-10-374-531-17
US-10-719-642-57
US-10-425-114-53273
US-10-719-642-11
US-10-767-701-45496
US-09-991-470-27
US-10-106-698-5606
US-09-764-864-1543
US-10-023-634-8

ALIGNMENTS

RESULT 1
US-09-798-689-34
; Sequence 34, Application US/09798689
; Publication No. US20030103973A1
; GENERAL INFORMATION:
; APPLICANT: Rockwell, Patricia
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: Method for Reducing Tumor Growth with VEGF Antagonists
; TITLE OF INVENTION: Combined With Radiation and Chemotherapy
; FILE REFERENCE: Sequence Listings 1-41 for 381-25 CIP
; CURRENT APPLICATION NUMBER: US/09/798,689
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/401,163
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/967,113
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 08/706,804
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/476,533
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/326,552
; PRIOR FILING DATE: 1994-10-20
; PRIOR APPLICATION NUMBER: 08/196,041
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentt in Ver. 2.1
; SEQ ID NO 34
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Mouse
US-09-798-689-34

Query Match 100.0%; Score 25; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
|||||

```

Db      1 ASESIS 6
RESULT 2
US-09-996-954B-10
; Sequence 10, Application US/09996954B
; Publication No. US20030157104A1
; GENERAL INFORMATION:
; APPLICANT: Waksal, Harlan W.
; TITLE OF INVENTION: Treatment of Refractory Human Tumors
; FILE REFERENCE: 11245-46605
; CURRENT APPLICATION NUMBER: US/09/996,954B
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/840,146
; PRIOR FILING DATE: 04-24-2001
; PRIOR APPLICATION NUMBER: 09/374,028
; PRIOR FILING DATE: 08-13-1999
; PRIOR APPLICATION NUMBER: 09/312,284
; PRIOR FILING DATE: 05-14-1999
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens-Rodent Chimera
US-09-996-954B-10
Query Match      100.0%; Score 25; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASESIS 6
Db      1 ASESIS 6
RESULT 3
US-10-239-656-30
; Sequence 30, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RIETHMULLER, GERT
; APPLICANT: LUTTERBUSE, RALF
; APPLICANT: BORSCHERT, KATRIN
; APPLICANT: KISCHEL, ROMAN
; APPLICANT: MAYER, MONIKA
; APPLICANT: HOFMEISTER, ROBERT
; TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
; FILE REFERENCE: 029976/0106
; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Anti-NKG2D
; OTHER INFORMATION: hybridoma 8G7C10 variable light chain CDR2
US-10-239-656-30
Query Match      100.0%; Score 25; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASESIS 6
Db      1 ASESIS 6
RESULT 4
US-10-239-656-40
; Sequence 40, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RIETHMULLER, GERT
; APPLICANT: LUTTERBUSE, RALF
; APPLICANT: BORSCHERT, KATRIN
; APPLICANT: KISCHEL, ROMAN
; APPLICANT: MAYER, MONIKA
; APPLICANT: HOFMEISTER, ROBERT
; TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
; TITLE OF INVENTION: TO AN EPITOPE OF THE NKG2D RECEPTOR COMPLEX
; FILE REFERENCE: 029976/0106
; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Anti-NKG2D
; OTHER INFORMATION: hybridoma 6ESA7 variable light chain CDR2
US-10-239-656-40
Query Match      100.0%; Score 25; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASESIS 6
Db      2 ASESIS 7
RESULT 5
US-10-374-600-101
; Sequence 101, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA: US/10/374,600
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C

```

```
; FILING DATE: 19-MAR-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-5288
; TELEFAX: (212) 425-7200
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-10-374-600-101.

Query Match      100.0%; Score 25; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASESIS 6
Db      2 ASESIS 7

RESULT 6
US-10-374-531-101
; Sequence 101, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,531
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-MAR-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-10-374-531-101

Query Match      100.0%; Score 25; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASESIS 6
Db      2 ASESIS 7

RESULT 7
US-10-029-386-30831
; Sequence 30831, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ACOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30831
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005547.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
US-10-029-386-30831

Query Match      100.0%; Score 25; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASESIS 6
Db      8 ASESIS 13

RESULT 8
US-10-424-599-191046
; Sequence 191046, Application US/10424599
; Publication No. US200400031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
```

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 191046
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Cloned Antibody CDR Sequences
US-10-424-599-191046

Query Match          100.0%; Score 25; DB 12; Length 46;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
Db 36 ASESIS 41

RESULT 9
US-10-379-151-3
; Sequence 3, Application US/10379151
; Publication No. US20040175692A1
; GENERAL INFORMATION:
; APPLICANT: Alexion Pharmaceuticals, Inc.
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: McWhirter, John
; TITLE OF INVENTION: CHRONIC LYMPHOCYTIC LEUKEMIA CELL LINE
; FILE REFERENCE: 60 CIP (1087-43 CIP)
; CURRENT APPLICATION NUMBER: US/10/379,151
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: PCT/US01/47931
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/254,113
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 61
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Cloned Antibody CDR Sequences
US-10-379-151-3

Query Match          100.0%; Score 25; DB 16; Length 61;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
Db 36 ASESIS 41

RESULT 10
US-10-379-151-18
; Sequence 18, Application US/10379151
; Publication No. US20040175692A1
; GENERAL INFORMATION:
; APPLICANT: Alexion Pharmaceuticals, Inc.
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: McWhirter, John
; TITLE OF INVENTION: CHRONIC LYMPHOCYTIC LEUKEMIA CELL LINE
; FILE REFERENCE: 60 CIP (1087-43 CIP)
; CURRENT APPLICATION NUMBER: US/10/379,151
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: PCT/US01/47931
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/254,113
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18

Query Match          100.0%; Score 25; DB 16; Length 61;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
Db 36 ASESIS 41

RESULT 11
US-10-424-599-229773
; Sequence 229773, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinhua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229773
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Cloned Antibody CDR Sequences
US-10-424-599-229773

Query Match          100.0%; Score 25; DB 12; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
Db 26 ASESIS 31

RESULT 12
US-09-991-470-25
; Sequence 25, Application US/09991470
; Patent No. US20020173477A1
; GENERAL INFORMATION:
; APPLICANT: Ruy S. Liou
; TITLE OF INVENTION: ANTI-IGE GENE THERAPY
; FILE REFERENCE: 99-5
; CURRENT APPLICATION NUMBER: US/09/991,470
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/397,569
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 107
; TYPE: PRT
; ORGANISM: human/murine
; OTHER INFORMATION: Cloned Antibody CDR Sequences
US-09-991-470-25

Query Match          100.0%; Score 25; DB 9; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
```

```
Db          51 ASESIS 56
|||||
RESULT 13
US-10-239-656-28
; Sequence 28, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RIETHMULLER, GERT
; APPLICANT: LUTTERBUSE, RALF
; APPLICANT: BORSCHERT, KATRIN
; APPLICANT: KISCHEL, ROMAN
; APPLICANT: MAYER, MONIKA
; APPLICANT: HOFMEISTER, ROBERT
; TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
; TITLE OF INVENTION: TO AN EPITOPE OF THE NKG2D RECEPTOR COMPLEX
; FILE REFERENCE: 029976/0106
; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Anti-NKG2D
; OTHER INFORMATION: hybridoma 8G7C10 variable light chain
US-10-239-656-28

Query Match          100.0%; Score 25; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 ASESIS 6
          |||||
Db          51 ASESIS 56

RESULT 14
US-10-239-656-38
; Sequence 38, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RIETHMULLER, GERT
; APPLICANT: LUTTERBUSE, RALF
; APPLICANT: BORSCHERT, KATRIN
; APPLICANT: KISCHEL, ROMAN
; APPLICANT: MAYER, MONIKA
; APPLICANT: HOFMEISTER, ROBERT
; TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
; TITLE OF INVENTION: TO AN EPITOPE OF THE NKG2D RECEPTOR COMPLEX
; FILE REFERENCE: 029976/0106
; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Anti-NKG2D
; OTHER INFORMATION: hybridoma 6B5A7 variable light chain
US-10-239-656-38

Query Match          100.0%; Score 25; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 ASESIS 6
          |||||
Db          51 ASESIS 56

US-10-239-656-38
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Anti-NKG2D
; OTHER INFORMATION: hybridoma 6B5A7 variable light chain
US-10-239-656-38

Query Match          100.0%; Score 25; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 ASESIS 6
          |||||
Db          51 ASESIS 56

US-10-453-698-129
; Sequence 129, Application US/10453698
; Publication No. US20040038308A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 82 CIP (1087-37 CIP)
; CURRENT APPLICATION NUMBER: US/10/453,698
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 129
; LENGTH: 107
; TYPE: PRT
; ORGANISM: human
US-10-453-698-129

Query Match          100.0%; Score 25; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 ASESIS 6
          |||||
Db          51 ASESIS 56

Search completed: October 6, 2004, 17:09:05
Job time : 37.3684 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 10.6316 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-10

Perfect score: 25

Sequence: 1 ASESIS 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/prodata/2/iaa/PTUS_COMB.pep:*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 25 | 100.0 | 7 | 2 | US-08-232-081B-5 |
| 2 | 25 | 100.0 | 7 | 2 | US-08-476-176B-54 |
| 3 | 25 | 100.0 | 7 | 3 | US-08-127-721A-54 |
| 4 | 25 | 100.0 | 7 | 3 | US-08-485-246A-54 |
| 5 | 25 | 100.0 | 99 | 3 | US-09-195-868-23 |
| 6 | 25 | 100.0 | 106 | 2 | US-08-800-198-4 |
| 7 | 25 | 100.0 | 106 | 3 | US-09-296-595-4 |
| 8 | 25 | 100.0 | 107 | 2 | US-08-232-081B-9 |
| 9 | 25 | 100.0 | 107 | 2 | US-08-476-176B-40 |
| 10 | 25 | 100.0 | 107 | 2 | US-08-476-176B-4 |
| 11 | 25 | 100.0 | 107 | 3 | US-08-127-721A-4 |
| 12 | 25 | 100.0 | 107 | 3 | US-08-485-246A-4 |
| 13 | 25 | 100.0 | 127 | 2 | US-08-476-176B-6 |
| 14 | 25 | 100.0 | 127 | 2 | US-08-476-176B-8 |
| 15 | 25 | 100.0 | 127 | 2 | US-08-476-176B-10 |
| 16 | 25 | 100.0 | 127 | 3 | US-08-127-721A-6 |
| 17 | 25 | 100.0 | 127 | 3 | US-08-127-721A-8 |
| 18 | 25 | 100.0 | 127 | 3 | US-08-127-721A-10 |
| 19 | 25 | 100.0 | 127 | 3 | US-08-485-246A-6 |
| 20 | 25 | 100.0 | 127 | 3 | US-08-485-246A-8 |
| 21 | 25 | 100.0 | 127 | 3 | US-08-485-246A-10 |
| 22 | 25 | 100.0 | 226 | 3 | US-09-195-868-30 |
| 23 | 25 | 100.0 | 239 | 2 | US-07-956-399-4 |
| 24 | 25 | 100.0 | 240 | 3 | US-08-800-198-8 |
| 25 | 25 | 100.0 | 240 | 3 | US-09-296-595-8 |
| 26 | 25 | 100.0 | 275 | 4 | US-09-510-322A-10 |
| 27 | 25 | 100.0 | 1189 | 3 | US-09-195-868-15 |

| | | | | | | |
|----|----|-------|------|---|---------------------|--------------------|
| 28 | 25 | 100.0 | 1229 | 3 | US-09-195-868-38 | Sequence 28, Appl |
| 29 | 24 | 96.0 | 109 | 1 | US-08-478-039-90 | Sequence 90, Appl |
| 30 | 24 | 96.0 | 109 | 1 | US-08-476-349A-90 | Sequence 90, Appl |
| 31 | 24 | 96.0 | 131 | 3 | US-08-589-399-3 | Sequence 3, Appl |
| 32 | 24 | 96.0 | 244 | 4 | US-08-244-369B-1 | Sequence 1, Appl |
| 33 | 24 | 96.0 | 371 | 4 | US-09-540-236-3776 | Sequence 3776, Ap |
| 34 | 24 | 96.0 | 1507 | 6 | 5268270-2 | Patent No. 5268270 |
| 35 | 23 | 92.0 | 240 | 2 | US-07-956-399-2 | Sequence 2, Appl |
| 36 | 23 | 92.0 | 337 | 4 | US-09-543-681A-6300 | Sequence 6300, Ap |
| 37 | 23 | 92.0 | 373 | 4 | US-09-066-281B-19 | Sequence 19, Appl |
| 38 | 23 | 92.0 | 373 | 4 | US-09-488-433C-19 | Sequence 19, Appl |
| 39 | 22 | 88.0 | 11 | 1 | US-07-942-245-518 | Sequence 518, App |
| 40 | 22 | 88.0 | 11 | 4 | US-09-339-922A-108 | Sequence 108, App |
| 41 | 22 | 88.0 | 11 | 4 | US-09-339-922A-110 | Sequence 110, App |
| 42 | 22 | 88.0 | 12 | 1 | US-08-221-580-1 | Sequence 1, Appl |
| 43 | 22 | 88.0 | 12 | 5 | PCT-US95-04018-63 | Sequence 63, Appl |
| 44 | 22 | 88.0 | 13 | 4 | US-09-339-922A-42 | Sequence 42, Appl |
| 45 | 22 | 88.0 | 13 | 4 | US-09-339-922A-82 | Sequence 82, Appl |

ALIGNMENTS

RESULT 1
US-08-232-081B-5
; Sequence 5, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:
; APPLICANT: NAKATANI, TOMOYUKI
; APPLICANT: GOMI, HIDEYUKI
; APPLICANT: WIJENES, JOHN
; APPLICANT: NOGUCHI, HIROSHI
; TITLE OF INVENTION: HUMANIZED B-B10
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.081B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 20-3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-232-081B-5

Query Match 100.0%; Score 25; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASESIS 6
Db 2 ASESIS 7

RESULT 2
US-08-476-176B-54
; Sequence 54, Application US/08476176B
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 5958708man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958708artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,176B
; FILING DATE: 27-SEPTEMBER-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 27-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5958708ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-176B-54
Query Match 100.0%; Score 25; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASESIS 6
Db 2 ASESIS 7
RESULT 3
US-08-127-721A-54
; Sequence 54, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066718artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover

STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,721A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6066718ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-127-721A-54
Query Match 100.0%; Score 25; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASESIS 6
Db 2 ASESIS 7
RESULT 4
US-08-485-246A-54
; Sequence 54, Application US/08485246A
; Patent No. 6072035
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6072035man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6072035artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,246A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6072035ak, Henry P.

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-296-595-4

Query Match 100.0%; Score 25; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 51 ASESIS 56

RESULT 8
US-08-232-081B-9
; Sequence 9, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:

APPLICANT: NAKATANI, TOMOYUKI
APPLICANT: GOMI, HIDEYUKI
APPLICANT: WIJDNES, JOHN
APPLICANT: NOGUCHI, HIROSHI
TITLE OF INVENTION: HUMANIZED B-B10
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,081B
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 20-3484
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-232-081B-9
Query Match 100.0%; Score 25; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 51 ASESIS 56

RESULT 9
US-08-232-081B-40
; Sequence 40, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:

APPLICANT: NAKATANI, TOMOYUKI
APPLICANT: GOMI, HIDEYUKI
APPLICANT: WIJDNES, JOHN
APPLICANT: NOGUCHI, HIROSHI
TITLE OF INVENTION: HUMANIZED B-B10
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,081B
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 20-3484
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-232-081B-40
Query Match 100.0%; Score 25; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 51 ASESIS 56

RESULT 10
US-08-476-176B-4
; Sequence 4, Application US/08476176B
; Patent No. 5958708
; GENERAL INFORMATION:

APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435

US-08-476-176B-4
Query Match 100.0%; Score 25; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 51 ASESIS 56

RESULT 10
US-08-476-176B-4
; Sequence 4, Application US/08476176B
; Patent No. 5958708
; GENERAL INFORMATION:

APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: Immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/127,721
;; FILING DATE: 27-SEPTEMBER-1993
;; APPLICATION NUMBER: US 07/952,802
;; FILING DATE: 25-SEPTEMBER-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5958708ak, Henry P.
;; REGISTRATION NUMBER: 33,200
;; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 277-5110
;; TELEFAX: (908) 277-4306
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 107 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-476-176B-4

Query Match 100.0%; Score 25; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 51 ASESIS 56

RESULT 11
US-08-127-721A-4
;; Sequence 4, Application US/08127721A
;; Patent No. 6066718
;; GENERAL INFORMATION:
;; APPLICANT: Hardman, No. 6066718man
;; APPLICANT: Kolbinger, Frank
;; APPLICANT: Saldanha, Jose
;; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
;; TITLE OF INVENTION: immunoglobulin isotype
;; NUMBER OF SEQUENCES: 55
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: No. 6066718artis Patent and Trademark Department
;; STREET: 59 Route 10
;; CITY: East Hanover
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07936-1090
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/127,721A
;; FILING DATE: 27-SEPTEMBER-1993
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/952,802
;; FILING DATE: 25-SEPTEMBER-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 6066718ak, Henry P.
;; REGISTRATION NUMBER: 33,200
;; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 277-5110
;; TELEFAX: (908) 277-4306
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 107 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-127-721A-4

Query Match 100.0%; Score 25; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASESIS 6
Db 51 ASESIS 56
RESULT 12
US-08-485-246A-4
;; Sequence 4, Application US/08485246A
;; Patent No. 6072035
;; GENERAL INFORMATION:
;; APPLICANT: Hardman, No. 6072035man
;; APPLICANT: Kolbinger, Frank
;; APPLICANT: Saldanha, Jose
;; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
;; TITLE OF INVENTION: immunoglobulin isotype
;; NUMBER OF SEQUENCES: 55
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: No. 6072035artis Patent Department
;; STREET: 59 Route 10
;; CITY: East Hanover
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07936-1080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/485,246A
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/127,721
;; FILING DATE: 27-SEPTEMBER-1993
;; APPLICATION NUMBER: US 07/952,802
;; FILING DATE: 25-SEPTEMBER-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 6072035ak, Henry P.
;; REGISTRATION NUMBER: 33,200
;; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908) 277-5110
;; TELEFAX: (908) 277-4306
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 107 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-485-246A-4

Query Match 100.0%; Score 25; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 51 ASESIS 56

RESULT 13
US-08-476-176B-6
;; Sequence 6, Application US/08476176B
;; Patent No. 5958708
;; GENERAL INFORMATION:
;; APPLICANT: Hardman, No. 5958708man
;; APPLICANT: Kolbinger, Frank

APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-176B-6

Query Match 100.0%; Score 25; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 71 ASESIS 76

RESULT 14
US-08-476-176B-8
Sequence 8, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-176B-8

Query Match 100.0%; Score 25; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 71 ASESIS 76

RESULT 15
US-08-476-176B-10
Sequence 10, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

```

;      LENGTH: 127 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-476-176B-10

Query Match      100.0%; Score 25; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ASESIS 6
       |||||
Db      71 ASESIS 76

Search completed: October 6, 2004, 16:36:33
Job time : 10.6316 secs

```

This Page Blank (uspio)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 30.3158 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-10

Perfect score: 25

Sequence: 1 ASESIS 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 25 | 100.0 | 6 | 3 AAY59315 | Aay59315 Light cha |
| 2 | 25 | 100.0 | 6 | 4 AAB37956 | Aab37956 Anti-EGFR |
| 3 | 25 | 100.0 | 6 | 5 AAU77788 | Aau77788 Mouse lig |
| 4 | 25 | 100.0 | 7 | 2 AAR37603 | Aar37603 hIL2R Ab |
| 5 | 25 | 100.0 | 7 | 2 AAY26991 | Aay26991 CDR2 doma |
| 6 | 25 | 100.0 | 7 | 5 AAU72844 | Aau72844 Anti-NKG2 |
| 7 | 25 | 100.0 | 7 | 5 AAU72852 | Aau72852 Anti-NKG2 |
| 8 | 25 | 100.0 | 61 | 5 ABB79660 | Abb79660 Chronic 1 |
| 9 | 25 | 100.0 | 61 | 5 ABB79661 | Abb79661 Chronic 1 |
| 10 | 25 | 100.0 | 85 | 3 AAG18882 | Aagi8882 Zea may |
| 11 | 25 | 100.0 | 89 | 3 AAG18881 | Aagi8881 Zea may |
| 12 | 25 | 100.0 | 104 | 2 AAW26795 | Aaw26795 Anti-gp34 |
| 13 | 25 | 100.0 | 106 | 2 AAW71241 | Aaw71241 Light cha |
| 14 | 25 | 100.0 | 107 | 2 AAR32129 | Aar32129 Anti-IL2R |
| 15 | 25 | 100.0 | 107 | 2 AAR37610 | Aar37610 B-B10 MAb |
| 16 | 25 | 100.0 | 107 | 2 AAR37612 | Aar37612 hIL2R Ab |
| 17 | 25 | 100.0 | 107 | 2 AAR50190 | Aar50190 Light cha |
| 18 | 25 | 100.0 | 107 | 2 AAW08949 | Aaw08949 Kappa lig |
| 19 | 25 | 100.0 | 107 | 2 AAW08948 | Aaw08948 Kappa lig |
| 20 | 25 | 100.0 | 107 | 2 AAY26979 | Aay26979 Light cha |
| 21 | 25 | 100.0 | 107 | 3 AAY70604 | Aay70604 VKappa re |
| 22 | 25 | 100.0 | 107 | 5 AAU72842 | Aau72842 Anti-NKG2 |
| 23 | 25 | 100.0 | 107 | 5 AAU72850 | Aau72850 Anti-NKG2 |
| 24 | 25 | 100.0 | 107 | 6 AAO16010 | Aao16010 Mouse int |
| 25 | 25 | 100.0 | 107 | 7 ABB80270 | Abb80270 Vitaxin 1 |

| | | | | | |
|----|----|-------|-----|------------|--------------------|
| 26 | 25 | 100.0 | 112 | 7 ADD13784 | Add13784 vkappa ex |
| 27 | 25 | 100.0 | 126 | 1 AAP90479 | Aap90479 Chimeric |
| 28 | 25 | 100.0 | 127 | 2 AAR50192 | Aar50192 Light cha |
| 29 | 25 | 100.0 | 127 | 2 AAR50187 | Aar50187 Light cha |
| 30 | 25 | 100.0 | 127 | 2 AAR50191 | Aar50191 Light cha |
| 31 | 25 | 100.0 | 127 | 2 AAW08945 | Aaw08945 Kappa lig |
| 32 | 25 | 100.0 | 127 | 2 AAW08941 | Aaw08941 Kappa lig |
| 33 | 25 | 100.0 | 127 | 2 AAW08946 | Aaw08946 Kappa lig |
| 34 | 25 | 100.0 | 127 | 2 AAW08943 | Aaw08943 Kappa lig |
| 35 | 25 | 100.0 | 127 | 2 AAY26981 | Aay26981 Light cha |
| 36 | 25 | 100.0 | 127 | 2 AAY26980 | Aay26980 Light cha |
| 37 | 25 | 100.0 | 127 | 2 AAY26982 | Aay26982 Light cha |
| 38 | 25 | 100.0 | 150 | 6 AAO30379 | Aao30379 Rabbit 20 |
| 39 | 25 | 100.0 | 214 | 6 ABP96757 | Abp96757 TSH recep |
| 40 | 25 | 100.0 | 214 | 6 ABP96766 | Abp96766 TSH recep |
| 41 | 25 | 100.0 | 214 | 6 ABP96758 | Abp96758 TSH recep |
| 42 | 25 | 100.0 | 214 | 6 ABP96765 | Abp96765 TSH recep |
| 43 | 25 | 100.0 | 219 | 7 ADC79230 | Adc79230 OC125-3.1 |
| 44 | 25 | 100.0 | 226 | 2 AAW26625 | Aaw26625 Signallin |
| 45 | 25 | 100.0 | 239 | 2 AAR34511 | Aar34511 FV(YU25). |

ALIGNMENTS

RESULT 1

AA59315

ID AAY59315 standard; peptide; 6 AA.

XX AC AAY59315;

XX DT 07-MAR-2000 (first entry)

XX DE Light chain hypervariable region, CDR2.

XX KW Hypervariable region; complementarity determining region; CDR; tumour;
single chain antibody; growth inhibitor; human; tumourigenesis; therapy;
protein receptor tyrosine kinase; light chain.

XX OS Mus sp.

XX PN WO9960023-A1.

XX PD 25-NOV-1999.

XX PF 14-MAY-1999; 99WO-US010741.

XX PR 15-MAY-1998; 98US-00079612.

XX PR 15-MAY-1998; 98US-0085613P.

XX PR 07-DEC-1998; 98US-00206138.

XX PA (IMCL-) INCLONE SYSTEMS INC.

XX PA (UABR-) UAB RES FOUND.

XX Waksal HW, Saleh MN, Robert F, Buchsbaum DJ;

XX WPI; 2000-062440/05.

XX N-PSDB; AAZ48629.

XX PT Treatment of human tumors, using a combination of radiation and a non-
radiolabeled protein receptor tyrosine kinase inhibitor.

XX PS Disclosure; Page 15; 31pp; English.

XX CC This sequence is the hypervariable region CDR2 (complementarity
determining region 2) of the light chain of a single chain antibody
derived from the murine antibody 225. The invention relates to a method
for inhibiting the growth of tumours in human patients by treating with
an effective amount of a combination of radiation and a non-radiolabelled
protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
which can lead to tumourigenesis. The method can be used in the treatment
of tumours of e.g. breast, lung, colon, kidney, bladder, head and neck,
ovary, prostate or brain. The administration of a suitable antibody to

CC the patient makes the tumour more susceptible to radiotherapy

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 25; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6

DB 1 ASESIS 6

RESULT 2

AAB37956

ID AAB37956 standard; protein; 6 AA.

XX AC AAB37956;

XX DT 12-VAR-2001 (first entry)

DE Anti-EGFR monoclonal antibody L chain V region CDR2 peptide sequence.

XX KW Refractory tumour growth inhibition; epidermal growth factor receptor;
KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
KW complementarity determining region; CDR.

XX OS Mus sp.

XX FN WO200069459-A1.

XX PD 23-NOV-2000.

XX PF 01-MAY-2000; 2000WO-US011756.

XX PR 14-MAY-1999; 99US-00312284.

XX PR 13-AUG-1999; 99US-00374028.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Waksal HW;

XX DR WPI; 2001-016160/02.

XX DR N-PSDB; AAC83239.

XX CC Epidermal growth factor receptor/human epidermal growth factor receptor-1
XX PT antagonist for inhibiting the growth of refractory tumors.
XX PS Disclosure; Page 14; 31pp; English.

XX CC This invention relates to a method for inhibiting the growth of
XX CC refractory tumors that are stimulated by a ligand of epidermal growth
XX CC factor receptor (EGFR) in human patients. The method involves treating
XX CC the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
XX CC optionally with a chemotherapeutic agent or radiation. The antagonist can
XX CC be for example a chimeric anti-EGFR monoclonal antibody, C225. The
XX CC EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
XX CC tumors such as tumors of breast, heart, lung, small intestine, colon,
XX CC spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas,
XX CC skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
XX CC liver, preferably squamous cell carcinomas. The present sequence
XX CC represents the light chain variable region complementarity determining
XX CC region 2 amino acid sequence of the chimeric anti-EGFR monoclonal
XX CC antibody C225 which is used in an example illustrating the method of the
XX CC invention

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 25; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6

DB 1 ASESIS 6

RESULT 3

AAU77788

ID AAU77788 standard; peptide; 6 AA.

XX AC AAU77788;

XX DT 05-JUN-2002 (first entry)

DE Mouse light chain hypervariable region (CDR2) of 225 antibody.

XX KW Mouse; light chain; antibody; hyperproliferative disease;
KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
KW seborrheic keratosis; warts; keloid scars; eczema; 255 antibody;
KW hypervariable region; CDR2; EGFR inhibitor.

XX OS Mus sp.

XX FN WO200211677-A2.

XX PD 14-FEB-2002.

XX PF 09-AUG-2001; 2001WO-US041647.

XX PR 09-AUG-2000; 2000US-00635974.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Teufel T;

XX DR WPI; 2002-257423/30.

XX DR N-PSDB; ABK11444.

XX CC Treating a mammal with hyperproliferative disease especially psoriasis,
XX PT stimulated by ligand of member of epidermal growth factor family of
XX PT receptors, by administering antagonist of the receptor.
XX PS Disclosure; Page 12; 28pp; English.

XX CC This invention relates to a novel method for treating a mammal with
XX CC hyperproliferative disease stimulated by a ligand of a member of the
XX CC epidermal growth factor (EGF) family of receptors. The method involves
XX CC administering an antibody or a defective receptor that is an antagonist
XX CC of a member of the EGF receptor family, or a combination of the
XX CC antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
XX CC The antibody used in the method of the invention acts as an epidermal
XX CC growth factor receptor (EGFR) antagonist by inhibiting EGFR /HER1
XX CC phosphorylation. The method of the invention is useful for treating a
XX CC mammal with hyperproliferative disease such as psoriasis, actinic
XX CC keratosis, seborrheic keratosis, warts, keloid scars and eczema
XX CC stimulated by a ligand of a member of the EGF family of receptor. This
XX CC sequence represents the murine anti-EGF 255 antibody light chain (LH)
XX CC hypervariable region (CDR2) used as an inhibitor of EGFR in the method of
XX CC the invention

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 25; DB 5; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6

DB 1 ASESIS 6

RESULT 4

AAR37603

ID AAR37603 standard; peptide; 7 AA.

XX

AC AAR37603;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-OCT-1993 (first entry)
 XX
 XX hIL2R Ab L chain V region CDR2.
 XX
 XX Complementarity-determining region; CDR; humanised; antibody; hIL2R;
 KW human; interleukin; IL-2; receptor; murine; anti-human; Ab; T-cell;
 KW monoclonal antibody; B-B10; mixed lymphocyte reaction; variable; V;
 KW region; PCR; framework; plasmid.
 XX
 OS Mus musculus.
 XX
 XX WO9311238-A1.
 XX
 XX 10-JUN-1993.
 XX
 XX 03-DEC-1992; 92WO-JP001583.
 XX
 XX 06-DEC-1991; 91JP-00323319.
 XX
 XX (SUMU) SUMITOMO PHARM CO LTD.
 PA (BIOT) BIOTEST PHARMA GMBH.
 PA (INNO-) INNOTHERAPIE LAB.
 XX
 XX Nakatani T, Gomi H, Wijdenes J, Noguchi H;
 XX
 XX WPI; 1993-197057/24.
 XX
 XX Humanised antibody comprising - CDR region of mouse MAB B-B10 specific
 PT for IL-2 receptor useful for treating carcinoma expressing IL-2 receptor.
 XX
 XX Claim 1; Page 43; 62pp; English.
 XX
 XX The sequences given in AAR37599-604 represent the complementarity-
 CC determining regions (CDRs) of a humanised antibody (Ab) which binds
 CC specifically to human interleukin (IL)-2 receptor (hIL2R). These CDRs
 CC were derived from the murine anti-human IL-2 receptor monoclonal Ab (MAB)
 CC B-B10 (see also AAQ43242-43). This MAB is antagonistic to the binding of
 CC IL-2 to the IL-2 receptor on human T-cells. It also inhibits the human
 CC mixed lymphocyte reaction. The cDNA encoding the variable (V) region of
 CC the B-B10 Ab was cloned by PCR and sequenced (see also AAQ43226-32 and
 CC AAQ43233-36). A human Ab with high levels of amino acid sequence homology
 CC to the murine sequence was selected and the framework of this Ab was
 CC bound with the B-B10 V region CDR and a part of the framework to design
 CC several kinds of the humanised B-B10 V region (see also AAQ43244-45). The
 CC DNA sequence coding this humanised B-B10 was synthesised and a plasmid
 CC expressing humanised B-B10 was constructed. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 XX Sequence 7 AA;
 XX
 XX Query Match 100.0%; Score 25; DB 2; Length 7;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 DB 2 ASESIS 7
 RESULT 5
 AAY26991
 ID AAY26991 standard; protein; 7 AA.
 XX
 XX AAY26991;
 AC
 XX
 XX 24-DEC-1999 (first entry)
 DT
 XX CDR2 domain reshaped human antibody light chain variable region.
 DE
 XX Diagnosis; IgE; immunoglobulin; body fluid; human; monoclonal antibody;
 XX

KW binding affinity; mouse; CDR; complementarity determining region;
 KW allergy.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX US5958708-A.
 PN
 XX 28-SEP-1999.
 PD
 XX 07-JUN-1995; 95US-00476176.
 PF
 XX 25-SEP-1992; 92US-00952802.
 PR
 XX 27-SEP-1993; 93US-00127721.
 XX
 XX (NOVS) NOVARTIS CORP.
 PA (TANO-) TANOX BIOSYSTEMS INC.
 XX
 XX Saldanha J, Kolbinger F, Hardman N;
 PI
 XX WPI; 1999-570765/48.
 DR
 XX new method for determining IgE levels in a sample.
 PT
 XX Claim 5; Col 67; 19pp; English.
 PS
 XX The invention relates to a method of determining IgE levels in a body
 CC fluid sample, by contacting the sample with a reshaped human monoclonal
 CC antibody (RA), having a binding affinity about equal to that of the murine
 CC CDR-donor antibody RES-C21 produced by the cell line 11133. The
 CC antibodies are useful in the diagnosis, prophylaxis and treatment of
 CC allergy. This sequence represents the complementarity determining region
 CC (CDR) 2 from the light chain variable domain of a reshaped human antibody
 XX
 XX Sequence 7 AA;
 SQ
 XX Query Match 100.0%; Score 25; DB 2; Length 7;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 DB 2 ASESIS 7
 RESULT 6
 AAU72844
 ID AAU72844 standard; peptide; 7 AA.
 XX
 XX AAU72844;
 AC
 XX
 XX 26-FEB-2002 (first entry)
 DT
 XX
 XX Anti-NKG2D hydridoma 8G7C10 variable light chain CDR3.
 DE
 XX Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;
 KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
 KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
 KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
 KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAP10;
 KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;
 KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;
 KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;
 KW p53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.
 XX
 XX Homo sapiens.
 OS
 XX WO200171005-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 26-MAR-2001; 2001WO-EP003414.
 XX
 XX

```

PR 24-MAR-2000; 2000EP-00106467.
XX
XX (KUFE/) KUPER P.
XX
PI Kufer P, Riethmueller G, Lutterbuese R, Borschert K, Kischel R;
PI Mayer M, Hofmeister R;
XX
XX MPI; 2002-055119/07.
XX
XX Multifunctional polypeptides comprising binding sites that specifically
PT recognize extracellular groups of the NKG2D receptor complex and domains
PT which function as receptors or ligands, useful for treating cancers and
PT infectious diseases.
XX
XX Example 3; Fig 16; 114pp; English.
XX
XX The invention relates to a multifunctional polypeptide comprising a
CC domain with a binding site that specifically recognises an extracellular
CC group of the NKG2D receptor complex and a second domain which functions
CC as a receptor or ligand. The polypeptide and its associated
CC polynucleotide are used for the preparation of a pharmaceutical
CC composition for the treatment of cancer, infections and/or autoimmune
CC conditions. The cancer may be a tumour of the head and neck, stomach,
CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,
CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,
CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.
CC The infectious diseases can be caused by viruses, bacteria, fungi,
CC protozoa or helminths. The autoimmune diseases include multiple
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and
CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D
CC receptor and the polypeptides of the invention
XX
XX Sequence 7 AA;
SQ
Query Match 100.0%; Score 25; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ASESIS 6
DB 2 ASESIS 7
|||||
RESULT 7
AAU72852
ID AAU72852 standard; peptide; 7 AA.
XX
XX AAU72852;
XX
XX 26-FEB-2002 (first entry)
XX
XX Anti-NKG2D hybridoma 6E5A7 variable light chain CDR3.
XX
XX Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;
XX autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;
XX intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;
XX prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;
XX sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAP10;
XX helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;
XX 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;
XX P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;
XX P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.
XX
XX Homo sapiens.
XX
XX WO200171005-A2.
XX
XX 27-SEP-2001.
XX
XX 26-MAR-2001; 2001WO-BF003414.
XX

```

```

PR 24-MAR-2000; 2000EP-00106467.
XX
XX (KUFE/) KUPER P.
XX
PI Kufer P, Riethmueller G, Lutterbuese R, Borschert K, Kischel R;
PI Mayer M, Hofmeister R;
XX
XX MPI; 2002-055119/07.
XX
XX Multifunctional polypeptides comprising binding sites that specifically
PT recognize extracellular groups of the NKG2D receptor complex and domains
PT which function as receptors or ligands, useful for treating cancers and
PT infectious diseases.
XX
XX Example 3; Fig 16; 114pp; English.
XX
XX The invention relates to a multifunctional polypeptide comprising a
CC domain with a binding site that specifically recognises an extracellular
CC group of the NKG2D receptor complex and a second domain which functions
CC as a receptor or ligand. The polypeptide and its associated
CC polynucleotide are used for the preparation of a pharmaceutical
CC composition for the treatment of cancer, infections and/or autoimmune
CC conditions. The cancer may be a tumour of the head and neck, stomach,
CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,
CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,
CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.
CC The infectious diseases can be caused by viruses, bacteria, fungi,
CC protozoa or helminths. The autoimmune diseases include multiple
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and
CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D
CC receptor and the polypeptides of the invention
XX
XX Sequence 7 AA;
SQ
Query Match 100.0%; Score 25; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ASESIS 6
DB 2 ASESIS 7
|||||
RESULT 8
ABB79660
ID ABB79660 standard; protein; 61 AA.
XX
XX ABB79660;
XX
XX 21-OCT-2002 (first entry)
XX
XX Chronic lymphocyte leukaemia specific scFv E1c CDR sequences.
XX
XX Chronic lymphocytic leukaemia; CLL; scFv; antibody; rabbit;
XX complementarity determining region; CDR; diagnosis; therapy.
XX
XX Oryctolagus cuniculus.
XX
XX Key Location/Qualifiers
XX Region 1..11 /label= LC-CDR1
XX /note= "light chain complementarity determining region 1"
XX Region 12..18 /label= LC-CDR2
XX /note= "light chain complementarity determining region 2"
XX Region 19..28 /label= LC-CDR3
XX /note= "light chain complementarity determining region 3"
XX Region 29..33 /label= HC-CDR1
XX /note= "heavy chain complementarity determining region 1"
XX

```

FT Region 34..48
 FT /label= HC-CDR2
 FT /note= "heavy chain complementarity determining region 2"
 FT Region 49..61
 FT /label= HC-CDR3
 FT /note= "heavy chain complementarity determining region 3"
 FT XX WO200259280-A2.
 FT PN 01-AUG-2002.
 FT XX 10-DEC-2001; 2001WO-US047931.
 FT XX 08-DEC-2000; 2000US-0254113P.
 FT XX (ALEX-) ALEXION PHARM INC.
 FT XX Bowdish KS, McWhirter J;
 FT XX WPI; 2002-599775/64.
 FT XX New chronic lymphocytic leukemia cell line (designated CLL-AAT), useful
 FT for studying, diagnosing or treating chronic lymphocytic leukemia (CLL)
 FT disease, or for identifying agents that are useful in the therapy of CLL
 FT disease.
 FT XX Claim 12; Fig 9B; 35pp; English.
 FT XX The present sequence comprises a summary of the complementarity
 CC determining regions (CDRs) of the light chain and heavy chain of the
 CC chronic lymphocytic leukaemia (CLL) specific rabbit scFv antibody E1c.
 CC Antibody regions separating the CDRs in the scFv are not given in the
 CC sequence. Rabbit scFv antibodies (see ABB79657-81) for B-CLL-specific
 CC cell surface antigens were selected using antibody phage display and cell
 CC surface panning. The invention provides a CLL line, CLL-AAT, derived from
 CC a B-CLL primary line without immortalisation by Epstein-Barr virus. The
 CC cell line is used to generate antibodies useful in the diagnosis and/or
 CC treatment of CLL. Antibodies derived from recombinant libraries may be
 CC selected using CLL-AAT as bait to isolate antibodies on the basis of
 CC specificity. Single chain antibodies are of particular use as they remain
 CC stable in the cytoplasm and retain intracellular binding activity. The
 CC binding of the present scFv antibody to primary human cells and cell
 CC lines was determined by whole cell ELISA as follows: CLL (primary tumours
 CC and CLL-AAT cell line) ++; normal, primary human B lymphocytes, +; non-
 CC Hodgkin's lymphoma cell line RL, -; Burkitt's lymphoma cell line Ramos, -
 CC ; and human erythroleukaemia cell line TF-1, -. A short linker separates
 CC the VL and VH regions of the scFv. The antigen recognised by this scFv
 CC was CD23, determined by immunoprecipitation and mass spectrometry
 XX Sequence 61 AA;
 SQ Query Match 100.0%; Score 25; DB 5; Length 61;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 DB 2 ASESIS 7
 RESULT 9
 ABB79661
 ID ABB79661 standard; protein; 61 AA.
 XX AC ABB79661;
 XX 21-OCT-2002 (first entry)
 DT Chronic lymphocyte leukaemia specific scFv F2d CDR sequences.
 DE Chronic lymphocytic leukaemia; CLL; scFv; antibody; rabbit;
 XX complementarity determining region; CDR; diagnosis; therapy.
 KW
 XX

OS Oryctolagus cuniculus.
 XX Key Location/Qualifiers
 FH Region 1..11
 FT /label= LC-CDR1
 FT /note= "light chain complementarity determining region 1"
 FT Region 12..18
 FT /label= LC-CDR2
 FT /note= "light chain complementarity determining region 2"
 FT Region 19..28
 FT /label= LC-CDR3
 FT /note= "light chain complementarity determining region 3"
 FT Region 29..33
 FT /label= HC-CDR1
 FT /note= "heavy chain complementarity determining region 1"
 FT Region 34..48
 FT /label= HC-CDR2
 FT /note= "heavy chain complementarity determining region 2"
 FT Region 49..61
 FT /label= HC-CDR3
 FT /note= "heavy chain complementarity determining region 3"
 FT XX WO200259280-A2.
 FT PN 01-AUG-2002.
 FT XX 10-DEC-2001; 2001WO-US047931.
 FT XX 08-DEC-2000; 2000US-0254113P.
 FT XX (ALEX-) ALEXION PHARM INC.
 FT XX Bowdish KS, McWhirter J;
 FT XX WPI; 2002-599775/64.
 FT XX New chronic lymphocytic leukemia cell line (designated CLL-AAT), useful
 FT for studying, diagnosing or treating chronic lymphocytic leukemia (CLL)
 FT disease, or for identifying agents that are useful in the therapy of CLL
 FT disease.
 FT XX Claim 12; Fig 9B; 35pp; English.
 CC The present sequence comprises a summary of the complementarity
 CC determining regions (CDRs) of the light chain and heavy chain of the
 CC chronic lymphocytic leukaemia (CLL) specific rabbit scFv antibody F2d.
 CC Antibody regions separating the CDRs in the scFv are not given in the
 CC sequence. Rabbit scFv antibodies (see ABB79657-81) for B-CLL-specific
 CC cell surface antigens were selected using antibody phage display and cell
 CC surface panning. The invention provides a CLL line, CLL-AAT, derived from
 CC a B-CLL primary line without immortalisation by Epstein-Barr virus. The
 CC cell line is used to generate antibodies useful in the diagnosis and/or
 CC treatment of CLL. Antibodies derived from recombinant libraries may be
 CC selected using CLL-AAT as bait to isolate antibodies on the basis of
 CC specificity. Single chain antibodies are of particular use as they remain
 CC stable in the cytoplasm and retain intracellular binding activity. The
 CC binding of the present scFv antibody to primary human cells and cell
 CC lines was determined by whole cell ELISA as follows: CLL (primary tumours
 CC and CLL-AAT cell line) ++; normal, primary human B lymphocytes, +; non-
 CC Hodgkin's lymphoma cell line RL, -; Burkitt's lymphoma cell line Ramos, -
 CC ; and human erythroleukaemia cell line TF-1, -. A short linker separates
 CC the VL and VH regions of the scFv. The antigen recognised by this scFv
 CC was CD23, determined by immunoprecipitation and mass spectrometry
 XX Sequence 61 AA;
 SQ Query Match 100.0%; Score 25; DB 5; Length 61;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 DB 2 ASESIS 7

us-09-635-974a-10.rag

Thu Oct 14 09:36:51 2004

```

RESULT 10
AAG18882
ID AAG18882 standard; protein; 85 AA.
XX
AC AAG18882;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 20468.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0124825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132464P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141342P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142230P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

```

| | | | | | |
|----|--------------|----------------|----|-------------|--|
| PR | 20-AUG-1999; | 99US-0149723P. | DT | 17-OCT-2000 | (first entry) |
| PR | 20-AUG-1999; | 99US-0149929P. | XX | | |
| PR | 23-AUG-1999; | 99US-0149902P. | DE | | Zea mays protein fragment SEQ ID NO: 20467. |
| PR | 23-AUG-1999; | 99US-0149930P. | XX | | |
| PR | 25-AUG-1999; | 99US-0150566P. | KW | | Protein identification; signal transduction pathway; metabolic pathway; |
| PR | 26-AUG-1999; | 99US-0150884P. | KW | | hybridisation assay; Genetic mapping; gene expression control; promoter; |
| PR | 27-AUG-1999; | 99US-0151065P. | KW | | termination sequence; corn. |
| PR | 27-AUG-1999; | 99US-0151066P. | XX | | |
| PR | 27-AUG-1999; | 99US-0151080P. | OS | | Zea mays subsp. mays. |
| PR | 30-AUG-1999; | 99US-0151303P. | XX | | |
| PR | 31-AUG-1999; | 99US-0151438P. | XX | | |
| PR | 01-SEP-1999; | 99US-0151930P. | PN | | EP1033405-A2. |
| PR | 07-SEP-1999; | 99US-0152363P. | XX | | |
| PR | 10-SEP-1999; | 99US-0153070P. | PD | | 06-SEP-2000. |
| PR | 13-SEP-1999; | 99US-0153758P. | XX | | |
| PR | 15-SEP-1999; | 99US-0154018P. | PF | | 25-FEB-2000; 2000EP-00301439. |
| PR | 16-SEP-1999; | 99US-0154039P. | XX | | |
| PR | 20-SEP-1999; | 99US-0154779P. | PR | | 25-FEB-1999; 99US-0121825P. |
| PR | 22-SEP-1999; | 99US-0155113P. | PR | | 05-MAR-1999; 99US-0123180P. |
| PR | 23-SEP-1999; | 99US-0155486P. | PR | | 09-MAR-1999; 99US-0123548P. |
| PR | 24-SEP-1999; | 99US-0155659P. | PR | | 23-MAR-1999; 99US-0125788P. |
| PR | 28-SEP-1999; | 99US-0156458P. | PR | | 25-MAR-1999; 99US-0126264P. |
| PR | 29-SEP-1999; | 99US-0156596P. | PR | | 29-MAR-1999; 99US-0126785P. |
| PR | 04-OCT-1999; | 99US-0157117P. | PR | | 01-APR-1999; 99US-0127462P. |
| PR | 05-OCT-1999; | 99US-0157533P. | PR | | 06-APR-1999; 99US-0128234P. |
| PR | 08-OCT-1999; | 99US-0158029P. | PR | | 08-APR-1999; 99US-0128714P. |
| PR | 08-OCT-1999; | 99US-0158232P. | PR | | 16-APR-1999; 99US-0129845P. |
| PR | 12-OCT-1999; | 99US-0158369P. | PR | | 19-APR-1999; 99US-0130077P. |
| PR | 13-OCT-1999; | 99US-0159293P. | PR | | 21-APR-1999; 99US-0130449P. |
| PR | 13-OCT-1999; | 99US-0159394P. | PR | | 23-APR-1999; 99US-0130510P. |
| PR | 13-OCT-1999; | 99US-0159395P. | PR | | 23-APR-1999; 99US-0130891P. |
| PR | 14-OCT-1999; | 99US-0159330P. | PR | | 28-APR-1999; 99US-0131449P. |
| PR | 14-OCT-1999; | 99US-0159331P. | PR | | 30-APR-1999; 99US-0132048P. |
| PR | 14-OCT-1999; | 99US-0159637P. | PR | | 30-APR-1999; 99US-0132407P. |
| PR | 14-OCT-1999; | 99US-0159638P. | PR | | 05-MAY-1999; 99US-0132484P. |
| PR | 18-OCT-1999; | 99US-0159584P. | PR | | 06-MAY-1999; 99US-0132485P. |
| PR | 21-OCT-1999; | 99US-0160741P. | PR | | 07-MAY-1999; 99US-0132487P. |
| PR | 21-OCT-1999; | 99US-0160767P. | PR | | 07-MAY-1999; 99US-0132863P. |
| PR | 21-OCT-1999; | 99US-0160768P. | PR | | 11-MAY-1999; 99US-0134256P. |
| PR | 21-OCT-1999; | 99US-0160770P. | PR | | 14-MAY-1999; 99US-0134218P. |
| PR | 21-OCT-1999; | 99US-0160814P. | PR | | 14-MAY-1999; 99US-0134219P. |
| PR | 21-OCT-1999; | 99US-0160815P. | PR | | 14-MAY-1999; 99US-0134221P. |
| PR | 22-OCT-1999; | 99US-0160800P. | PR | | 14-MAY-1999; 99US-0134370P. |
| PR | 22-OCT-1999; | 99US-0160981P. | PR | | 18-MAY-1999; 99US-0134768P. |
| PR | 22-OCT-1999; | 99US-0160989P. | PR | | 19-MAY-1999; 99US-0134941P. |
| PR | 25-OCT-1999; | 99US-0161404P. | PR | | 20-MAY-1999; 99US-0135124P. |
| PR | 25-OCT-1999; | 99US-0161405P. | PR | | 21-MAY-1999; 99US-0135353P. |
| PR | 25-OCT-1999; | 99US-0161406P. | PR | | 24-MAY-1999; 99US-0135629P. |
| PR | 26-OCT-1999; | 99US-0161359P. | PR | | 25-MAY-1999; 99US-0136021P. |
| PR | 26-OCT-1999; | 99US-0161360P. | PR | | 27-MAY-1999; 99US-0136382P. |
| PR | 28-OCT-1999; | 99US-0161361P. | PR | | 28-MAY-1999; 99US-0136782P. |
| PR | 28-OCT-1999; | 99US-0161320P. | PR | | 01-JUN-1999; 99US-0137222P. |
| PR | 28-OCT-1999; | 99US-0161992P. | PR | | 03-JUN-1999; 99US-0137528P. |
| PR | 28-OCT-1999; | 99US-0161993P. | PR | | 04-JUN-1999; 99US-0137502P. |
| PR | 29-OCT-1999; | 99US-0162142P. | PR | | 07-JUN-1999; 99US-0137724P. |
| PR | | | PR | | 08-JUN-1999; 99US-0138094P. |
| PR | | | PR | | 10-JUN-1999; 99US-0138540P. |
| PR | | | PR | | 10-JUN-1999; 99US-0138847P. |
| PR | | | PR | | 14-JUN-1999; 99US-0139119P. |
| PR | | | PR | | 16-JUN-1999; 99US-0139452P. |
| PR | | | PR | | 16-JUN-1999; 99US-0139453P. |
| PR | | | PR | | 17-JUN-1999; 99US-0139492P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139454P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139455P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139456P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139457P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139458P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139459P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139460P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139461P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139462P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139463P. |
| PR | | | PR | | 18-JUN-1999; 99US-0139750P. |

Query Match Best Local Similarity 100.0%; Score 25; DB 3; Length 85; Mismatches 6; Conservative 0; Indels 0; Gaps 0;

QY 1 AGESIS 6
Db 22 AGESIS 27

RESULT 11
AAG18881
ID AAG18881 standard; protein; 89 AA.
XX
AC AAG18881;
XX

PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140994P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151103P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 25; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASESIS 6
Db 26 ASESIS 31

RESULT 12
AAW26795
ID AAW26795 standard; protein; 104 AA.
XX AAW26795;
AC AAW26795;
XX 14-SEP-1998 (first entry)

DT Anti-gp54 MAb 48-127 light chain variable region.
DE Single chain antibody; scfv; monoclonal antibody; MAb; 48-127; trop-2;
KW gp54; antigen; cancer marker; bladder cancer; breast cancer;

KW cervix cancer; prostate cancer; metastasis; diagnosis; therapy;

```

KW immunotherapy.
XX Mus sp.
XX
XX Key Location/Qualifiers
XX Region 21..31
XX Region /label= CDR1
XX Region 47..54
XX Region /label= CDR2
XX Region 76..84
XX Region /label= CDR3
XX
XX WO9812227-A1.
XX
XX 26-MAR-1998.
XX
XX 19-SEP-1997; 97WO-CA000690.
XX
XX 19-SEP-1996; 96US-0026320P.
XX
XX (DIAG-) DIAGNOCURE INC.
XX
XX Fradet Y, Darveau A;
XX
XX WPI; 1998-217211/19.
XX
XX N-PSDB; AAV26766.
XX
XX New recombinant antibodies to gp54 antigen - used for the detection of
XX cancers and metastases and for therapy of cancers expressing the gp54
XX antigen.
XX
XX Claim 1; Fig 2; 33pp; English.
XX
XX This polypeptide comprises the light chain variable region of murine
XX monoclonal antibody (Mab) 48-127. This Mab recognises an epitope of gp54
XX antigen (trop-2), a cancer marker expressed at the surface of bladder,
XX breast, uterine and prostate cancers. Mab 48-127 (IgG1) reacts with a
XX majority of human bladder cell lines and with human breast carcinoma cell
XX lines. The invention relates to novel single chain antibody (scFv)
XX molecules, including a 48-127 scFv (see AAW26799) that comprises the 48-
XX 127 VH and VL chains joined via a linker. Such scFvs can be used to
XX detect cancers and micrometastases, e.g. bladder, prostate, breast or
XX uterine cervix cancers, and may be linked to specific toxins for use in
XX cancer therapy. The scFvs are not immunogenic and are distributed and
XX captured by gp45-bearing target cells more quickly than conventional
XX antibodies
XX
XX Sequence 104 AA;
XX
XX Query Match 100.0%; Score 25; DB 2; Length 104;
XX Best Local Similarity 100.0%; Pred. No. 99;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ASESIS 6
XX |||||
XX 48 ASESIS 53
XX
XX Db
XX
XX RESULT 13
XX AAW71241
XX ID AAW71241 standard; protein; 106 AA.
XX
XX AC AAW71241;
XX
XX DT 20-NOV-1998 (first entry)
XX
XX DE Light chain variable domain of antibody 14E1.
XX
XX Light chain variable domain; antibody 14E1;
XX human epidermal growth factor receptor; EGF-R; treatment; carcinoma;
XX EGF-RvIII; in vitro purging; breast carcinoma; ss.
XX
XX Mus sp.
XX
XX WO9836074-A2.
XX
XX 20-AUG-1998.
XX
XX 13-FEB-1998; 98WO-IB000413.
XX
XX 13-FEB-1997; 97US-00800198.
XX
XX (SCHD ) SCHERING AG.
XX
XX Wels WS, Schmidt M, Vakalopoulou E, Schneider DW;
XX
XX WPI; 1998-467176/40.
XX
XX N-PSDB; AAV54788.
XX
XX New single chain polypeptide that binds to epidermal growth factor
XX receptor - is derived from antibody 14E, used for treatment and diagnosis
XX of cancer or other cell proliferative diseases.
XX
XX Example 1; Fig 11D; 48pp; English.
XX
XX The present sequence represents the light chain variable domain of murine
XX monoclonal antibody 14E1. This antibody is active against the human
XX epidermal growth factor receptor (EGF-R). The sequence is used to
XX construct a single-chain polypeptide (scFv) that has binding affinity for
XX an EGF-R. The scFv comprises two segments which contain the binding parts
XX of the heavy and light chain variable domains of the monoclonal antibody
XX 14E1. The scFvs of the invention are used to treat carcinomas (gliomas,
XX melanomas and tumours) that overexpress EGF-R, specifically EGF-RvIII,
XX relative to normal cells, particularly by blocking proliferation of such
XX cells by preventing receptor activation. They can also be used for in
XX vitro purging. When the scFv is appropriately labelled, it can be used
XX diagnostically to image cells that express EGF-R, specifically breast
XX carcinoma
XX
XX Sequence 106 AA;
XX
XX Query Match 100.0%; Score 25; DB 2; Length 106;
XX Best Local Similarity 100.0%; Pred. No. 1e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ASESIS 6
XX |||||
XX 51 ASESIS 56
XX
XX Db
XX
XX RESULT 14
XX AAR32129
XX ID AAR32129 standard; protein; 107 AA.
XX
XX AC AAR32129;
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 10-MAR-2003 (revised)
XX
XX DT 02-JUN-1993 (first entry)
XX
XX DE Anti-IL2R beta antibody Mab A41 light chain variable region.
XX
XX immunosuppression; tissue transplantation; graft; L chain; V region;
XX T-helper cell inhibition; transplant rejection; MAb;
XX interleukin-2 receptor.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..96
XX Region /label= V-region
XX Region 97..107
XX Region /label= J-region
XX
XX DE4143214-A1.
XX
XX

```

PD 28-JAN-1993.
 XX
 PF 30-DEC-1991; 91DE-04143214.
 XX
 PR 25-JUL-1991; 91DE-04124759.
 XX
 PA (BOE) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Weidle U, Scheuer W, Kaluza B, Riethmueller G;
 XX
 DR WPI; 1993-037582/05.
 DR N-PSDB; AAQ36615.
 XX
 PT Synergistic antibody compsn. for use as immunosuppressant - comprises
 PT monoclonal anti-CD4 antibodies and monoclonal anti-IL2R alpha- or anti-
 PT IL2R beta antibodies.
 XX
 PS Claim 8; Page 17; 18pp; German.
 XX
 CC This sequence is the light chain variable region of a preferred anti-IL2R
 CC beta monoclonal antibody for use in the claimed synergistic composition.
 CC The anti-IL2R beta antibody is used with at least one anti-CD4 antibody.
 CC Individually the antibodies are strongly inhibiting and when used
 CC together their immunosuppressive properties are improved they
 CC synergistically inhibit T-helper cell proliferation to effectively
 CC inhibit transplant rejection at low doses without significantly reducing
 CC the general immune response. See AAQ36607-Q36616. (Updated on 10-MAR-2003
 CC to add missing OS field.) (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 107 AA;
 Query Match 100.0%; Score 25; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 Db 51 ASESIS 56
 RESULT 15
 AAR37610
 ID AAR37610 standard; protein; 107 AA.
 XX
 AC AAR37610;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-OCT-1993 (first entry)
 XX
 DE B-B10 MAb L chain V region.
 XX
 KW Complementarity-determining region; CDR; humanised; antibody; hIL2R;
 KW human; interleukin; IL-2; receptor; murine; anti-human; Ab; T-cell;
 KW monoclonal antibody; B-B10; mixed lymphocyte reaction; variable; V;
 KW region; PCR; framework; plasmid; heavy; H; light; L.
 XX
 OS Mus musculus.
 XX
 FN WO9311238-A1.
 XX
 PD 10-JUN-1993.
 XX
 PF 03-DEC-1992; 92WO-JP001583.
 XX
 PR 06-DEC-1991; 91JP-00323319.
 XX
 PA (SUMU) SUMITOMO PHARM CO LTD.
 PA (BIOT) BIOTEST PHARMA GMBH.
 PA (INNO-) INNOTHERAPIE LAB.
 XX
 PI Nakatani T, Gomi H, Wijdenes J, Noguchi H;
 XX
 DR WPI; 1993-197057/24.

DR N-PSDB; AAQ43243.
 XX
 PT Humanised antibody comprising - CDR region of mouse MAB B-B10 specific
 PT for IL-2 receptor useful for treating carcinoma expressing IL-2 receptor.
 XX
 PS Disclosure; Fig 2; 62pp; English.
 XX
 CC The sequences given in AAR37609-10 represent the heavy (H) and light (L)
 CC chain variable (V) regions of the murine anti-human IL-2 receptor
 CC monoclonal antibody (MAB) B-B10, respectively. This MAB was used in the
 CC construction of a humanised antibody (Ab) which binds specifically to
 CC human interleukin (IL)-2 receptor (hIL2R). The complementarity-
 CC determining regions (CDRs) for the hIL2R MAB were derived from B-B10 (see
 CC also AAR37599-04). The hIL2R MAB is antagonistic to the binding of IL-2
 CC to the IL-2 receptor on human T-cells. It also inhibits the human mixed
 CC lymphocyte reaction. The cDNA encoding the variable (V) region of the B-
 CC B10 Ab was cloned by PCR and sequenced (see also AAQ43226-32 and AAQ43233
 CC -36). A human Ab with high levels of amino acid sequence homology to the
 CC murine sequence was selected and the framework of this Ab was bound with
 CC the B-B10 V region CDR and a part of the framework to design several
 CC kinds of the humanised B-B10 V region. The DNA sequence coding this
 CC humanised B-B10 was synthesised and a plasmid expressing humanised B-B10
 CC was constructed. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 107 AA;
 Query Match 100.0%; Score 25; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASESIS 6
 Db 51 ASESIS 56
 Search completed: October 6, 2004, 16:29:48
 Job time : 32.3158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 3.68421 Seconds
(without alignments)
84.800 Million cell updates/sec

Title: US-09-635-974A-10
Perfect score: 25
Sequence: 1 ASESIS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 25 | 100.0 | 92 | 1 | P01690 oryctolagus |
| 2 | 25 | 100.0 | 115 | 1 | P01642 mus musculus |
| 3 | 25 | 100.0 | 368 | 1 | P26213 aspergillus |
| 4 | 25 | 100.0 | 465 | 1 | Q9ewh8 streptomyces |
| 5 | 25 | 100.0 | 451 | 1 | Q82877 streptomyces |
| 6 | 25 | 100.0 | 828 | 1 | Q96352 homo sapien |
| 7 | 25 | 100.0 | 1427 | 1 | P30622 homo sapien |
| 8 | 24 | 96.0 | 491 | 1 | P40499 saccharomyc |
| 9 | 24 | 96.0 | 635 | 1 | Q9pif9 chlamydia m |
| 10 | 24 | 96.0 | 893 | 1 | Q24738 drosophila |
| 11 | 24 | 96.0 | 1532 | 1 | P09790 neisseria g |
| 12 | 23 | 92.0 | 316 | 1 | Q84010 chlamydia t |
| 13 | 23 | 92.0 | 373 | 1 | Q9ubf1 homo sapien |
| 14 | 23 | 92.0 | 590 | 1 | Q9chl8 lactococcus |
| 15 | 23 | 92.0 | 590 | 1 | P97046 lactococcus |
| 16 | 23 | 92.0 | 616 | 1 | Q8z4n2 salmoneilla |
| 17 | 23 | 92.0 | 616 | 1 | Q8z4n2 salmoneilla |
| 18 | 23 | 92.0 | 711 | 1 | P35819 neisseria g |
| 19 | 22 | 88.0 | 108 | 1 | P01600 homo sapien |
| 20 | 22 | 88.0 | 109 | 1 | P01624 homo sapien |
| 21 | 22 | 88.0 | 117 | 1 | P01691 oryctolagus |
| 22 | 22 | 88.0 | 117 | 1 | P01602 homo sapien |
| 23 | 22 | 88.0 | 129 | 1 | P04431 homo sapien |
| 24 | 22 | 88.0 | 148 | 1 | Q83mw0 tropheryma |
| 25 | 22 | 88.0 | 185 | 1 | Q08793 bacillus su |
| 26 | 22 | 88.0 | 188 | 1 | P36303 kennedy's ye |
| 27 | 22 | 88.0 | 204 | 1 | P57893 pasteurella |
| 28 | 22 | 88.0 | 260 | 1 | P13634 mus musculus |
| 29 | 22 | 88.0 | 264 | 1 | Q97mks plasmodium |
| 30 | 22 | 88.0 | 283 | 1 | Y191_CLOAB |
| 31 | 22 | 88.0 | 324 | 1 | Y191_CLOAB |
| 32 | 22 | 88.0 | 331 | 1 | Y191_CLOAB |
| 33 | 22 | 88.0 | 337 | 1 | ASPC_AGICL |

| | | | | | | |
|----|----|------|-----|---|------------|---------------------|
| 34 | 22 | 88.0 | 362 | 1 | ASPG_PSEEL | O68897 pseudomonas |
| 35 | 22 | 88.0 | 397 | 1 | GDN_MOUSE | Q07235 mus musculus |
| 36 | 22 | 88.0 | 414 | 1 | YH42_MYCTU | Q10523 mycobacteri |
| 37 | 22 | 88.0 | 418 | 1 | YCCB_ECOLI | P36661 escherichia |
| 38 | 22 | 88.0 | 427 | 1 | ARP3_SCHPO | P32390 schizosacch |
| 39 | 22 | 88.0 | 432 | 1 | CLPX_CLOAB | Q97ft7 clostridium |
| 40 | 22 | 88.0 | 433 | 1 | HISX_ECO57 | Q8x8t3 escherichia |
| 41 | 22 | 88.0 | 433 | 1 | HISX_ECOL6 | O8f552 escherichia |
| 42 | 22 | 88.0 | 433 | 1 | HISX_ECOL6 | P05988 escherichia |
| 43 | 22 | 88.0 | 433 | 1 | HISX_SHIFL | P59401 shigella fl |
| 44 | 22 | 88.0 | 446 | 1 | NU4M_DROYA | P07707 drosophila |
| 45 | 22 | 88.0 | 452 | 1 | YEDV_ECOLI | P76339 escherichia |

ALIGNMENTS

RESULT 1

KV09_RABIT
ID KV09_RABIT STANDARD; PRT; 92 AA.
AC P01690;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V region 3381 (Fragment)
OS Oryctolagus cuniculus (Rabbit)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=75176905; PubMed=1094456;
RA Margolies M.N., Cannon L.E. III, Strosberg A.D., Haber E.;
RT "Diversity of light chain variable region sequences among rabbit
antibodies elicited by the same antigens."
RL Proc. Natl. Acad. Sci. U.S.A. 72:2180-2184 (1975).
CC -!- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM ANTIBODY TO TYPE III
CC PIR; A01953; KVRB38.
DR HSSP; P01607; IREI.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 >92
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 9730 MW; 9A36B15913CB3BEE CRC64;

Query Match 100.0%; Score 25; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 25 ASESIS 30

RESULT 2

KV51_MOUSE
ID KV51_MOUSE STANDARD; PRT; 115 AA.
AC P01642;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig kappa chain V-V region L7 precursor (Fragment).
OS Mus musculus (Mouse).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X58892; CAA41693.1; -

CC PIR; S17980; S17980.

CC InterPro; IPR000743; Glyco_hydro_28.

CC InterPro; IPR006626; PpH1.

CC Pfam; PF00395; Glyco_hydro_28; 1.

CC SMART; SM00710; PpH1; 6.

CC PROSITE; PS00502; POLYGALACTURONASE; 1.

CC Hydrolase; Glycosidase; Cell wall; Signal; Zymogen; Multigene family;

CC Glycoprotein.

CC SIGNAL 1 18 POTENTIAL.

CC FT PROPEP 19 31 POTENTIAL.

CC FT CHAIN 32 368 POLYGALACTURONASE 1.

CC FT ACT_SITE 229 229 PROBABLE.

CC FT CARBOHYD 246 246 N-LINKED (GLCNAC...) (PROBABLE).

CC SQ SEQUENCE 368 AA; 38108 MW; CD9B46A9A99B5102 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 368;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
| | | | |
DB 43 ASESIS 48

RESULT 4
ENG STRCO STANDARD; PRT; 465 AA.
AC QSEWNS;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE GRP-binding protein engA.
GN ENGA OR SC01758 OR 2SCI34.11C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: GTPase of unknown physiological role.
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. ENGA subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AL939110; CAC12931.1; ALT_INIT.
CC HAVAP; MF_00195; -; 1.

CC -----

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81220975; PubMed=6264318;
RA Pech M., Hochli J., Schnell H., Zachau H.G.;
RT "Differences between germ-line and rearranged immunoglobulin V kappa
RT coding sequences suggest a localized mutation mechanism.";
RL Nature 291:668-670(1981).
CC -!- MISCELLANEOUS: THERE APPEAR TO BE TWO POSSIBLE SPLICING JUNCTIONS AT
CC THE 3' END OF THE INTRON. THE ALTERNATE WOULD CODE FOR A PROTEIN
CC LACKING RESIDUES 17-19.

CC PIR; A01925; KMSL7.
CC PDB; 1UJO; 18-FEB-03.
CC PDB; 1UJP; 18-FEB-03.
CC PDB; 1UJX; 18-FEB-03.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC Immunoglobulin V region; Signal; 3D-structure.

CC SIGNAL 1 20
FT CHAIN 21 >115 IG KAPPA CHAIN V-V REGION L7.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 63 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 >115 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 115 115
FT SEQUENCE 115 AA; 12615 MW; C17BEC758C577E00 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
| | | | |
DB 71 ASESIS 76

RESULT 3
PG1L ASPNG STANDARD; PRT; 368 AA.
AC P262T3;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polyalacturonase I precursor (EC 3.2.1.15) (PG-I) (Pectinase).
GN PGAI
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-43; 107-117 AND 205-221.
RC STRAIN=CBS 120.49 / N400;
RA MEDLINE=92035101; PubMed=19341135;
RA Bussink H.J.D., Brouwer K., de Graaff L.H., Kester H.C.M., Visser J.;
RT "Identification and characterization of a second polyalacturonase
RT gene of Aspergillus niger.";
RL Curr. Genet. 20:301-307(1991).
CC -!- FUNCTION: Involved in maceration and soft-rotting of plant tissue.
CC -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -!- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP_OBG.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OGB.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00650; MG442; 2.
DR TIGRFAMs; TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 33 40 GTP 1 (POTENTIAL).
FT NP_BIND 80 84 GTP 1 (POTENTIAL).
FT NP_BIND 142 145 GTP 1 (POTENTIAL).
FT NP_BIND 208 215 GTP 2 (POTENTIAL).
FT NP_BIND 255 259 GTP 2 (POTENTIAL).
FT NP_BIND 320 323 GTP 2 (POTENTIAL).
SQ SEQUENCE 465 AA; 50344 MW; 9E053E7C00C859F8 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 293 ASESIS 298

RESULT 5
ENGA-STRAW STANDARD; PRT; 491 AA.
AC Q828Y7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE GTP-binding protein engA.
ENGA OR S46524.
GN Streptomyces avermitilis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22408306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -!- FUNCTION: GTPase of unknown physiological role.
CC -!- SIMILARITY: belongs to the era/trmE family of GTP-binding
CC proteins. Enga subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; AP005047; BAC74235.1; -.
DR HAMAP; MP_00195; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OGB.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00650; MG442; 2.
DR TIGRFAMs; TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 59 66 GTP 1 (POTENTIAL).
FT NP_BIND 106 110 GTP 1 (POTENTIAL).
FT NP_BIND 168 171 GTP 1 (POTENTIAL).
FT NP_BIND 234 241 GTP 2 (POTENTIAL).
FT NP_BIND 281 285 GTP 2 (POTENTIAL).
FT NP_BIND 346 349 GTP 2 (POTENTIAL).
SQ SEQUENCE 491 AA; 53353 MW; 7BADDBE049D14AC CRC64;

Query Match 100.0%; Score 25; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 319 ASESIS 324

RESULT 6
COG3 HUMAN STANDARD; PRT; 828 AA.
ID Q96J2; Q9BZ92;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Conserved oligomeric Golgi complex component 3 (Vesicle docking
DE protein SEC34 homolog) (p94).
GN COG3 OR SEC34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Fetal brain;
RX MEDLINE=21303583; PubMed=11292827;
RA Suvorova E.S., Kurten R.C., Lupashin V.V.;
RT "Identification of a human ortholog of Sec34p as a component of the
RT cis-Golgi vesicle tethering machinery.";
RL J. Biol. Chem. 276:22810-22818(2001).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=22050606; PubMed=11929878;
RA Loh E., Hong W.;
RT "Sec34 is implicated in traffic from the endoplasmic reticulum to the
RT Golgi and exists in a complex with GTC-90 and ldlbp.";
RL J. Biol. Chem. 277:21955-21961(2002).
CC -!- FUNCTION: Involved in ER-Golgi transport.
CC -!- SUBUNIT: Component of the conserved oligomeric Golgi complex which
CC is composed of eight different subunits and is required for normal
CC Golgi morphology and localization (By similarity).
CC -!- SUBCELLULAR LOCATION: Golgi. Associated with the peripheral
CC membrane of cis/medial cisternae.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest levels in
CC pancreas and testis and lowest levels in lung.
CC -!- SIMILARITY: Belongs to the COG3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF349676; AAK66974.1; -
 CC EMBL; AF332595; AAK06848.1; -
 CC Genew; HGNC:18619; COG3.
 CC MIM; 606975; -
 CC InterPro; IPR007265; Sec34.
 CC Pfam; PF04136; Sec34; 1.
 CC Transport; Protein transport; Golgi stack; Membrane.
 KW CONFLICT 106 107 HE -> QQ (IN REF. 2).
 FT CONFLICT 747 747 N -> S (IN REF. 2).
 FT CONFLICT 747 747 N -> S (IN REF. 2).
 SQ SEQUENCE 828 AA; 94079 MW; 8DFC7ED3F03B7265 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 828;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 DB 595 ASESIS 600

RESULT 7
 REST HUMAN
 ID REST HUMAN STANDARD; PRT; 1427 AA.
 AC P30622;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
 DE Sternberg intermediate filament associated protein).
 GN RSN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peripheral blood monocytes;
 RX MEDLINE=92289675; PubMed=1600942;
 RA Bilbe G., Delabie J., Brueggan J., Richner H., Asselbergs F.A.M.,
 RA Cerletti N., Sorg C., Odink K., Tarsay L., Wissendanger W.,
 RA de Wolf-Peters C., Shipman R.;
 RT "Restin: a novel intermediate filament-associated protein highly
 RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
 RL EMBO J. 11:2103-2113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92405160; PubMed=1356075;
 RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
 RT "CLIP-170 links endocytic vesicles to microtubules.";
 RL Cell 70:887-900(1992).
 CC -!- FUNCTION: SEEMS TO BE AN INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoID=P30622-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoID=P30622-2; Sequence=VSP 000765;
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
 CC OF HODGKIN'S DISEASE.
 CC -!- SIMILARITY: Contains 2 CAP-Gly domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X64838; CAA46050.1; -
 CC EMBL; M97501; AAA35693.1; -
 CC PIR; S22695; S22695.
 CC Genew; HGNC:10461; RSN.
 CC MIM; 179838; -
 CC GO; GO:0005768; C:Endosome; TAS.
 CC GO; GO:0005882; C:Intermediate filament; TAS.
 CC GO; GO:0015630; C:Microtubule cytoskeleton; TAS.
 CC GO; GO:0008017; F:Microtubule binding; TAS.
 CC GO; GO:0008699; F:nonselective vesicle transport; TAS.
 CC InterPro; IPR009338; CAP-Gly.
 CC InterPro; IPR001878; Znf CCHC.
 CC Pfam; PF01302; CAP GLY; 2.
 CC SMART; SM00343; Znf C2HC; 1.
 CC PROSITE; PS00845; CAP GLY 1; 2.
 CC PROSITE; PS0245; CAP GLY_2; 2.
 CC Cytoskeleton; Microtubule; Coiled coil; Repeat; Alternative splicing.
 KW DOMAIN 78 120 CAP-GLY 1.
 FT DOMAIN 143 204 SER-RICH.
 FT DOMAIN 232 274 CAP-GLY 2.
 FT DOMAIN 304 331 SER-RICH.
 FT DOMAIN 350 1342 COILED COIL (POTENTIAL).
 FT DOMAIN 1408 1421 CCHC-BOX.
 FT VARSPPLIC 457 491 Missing (in isoform Short).
 FT CONFLICT 1069 1069 /FTIG=VSP 000765.
 FT CONFLICT 1427 AA; 160989 MW; 0A4F166DD94254E8 CRC64;
 SQ SEQUENCE 1427 AA; 160989 MW; 0A4F166DD94254E8 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 1 ASESIS 6
 DB 192 ASESIS 197

RESULT 8
 YIJ0 YEAST
 ID YIJ0 YEAST STANDARD; PRT; 491 AA.
 AC P40459;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 56.3 kDa protein in SGAL-KTR7 intergenic region.
 GN YIJ090W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=97313266; PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Holsen T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Moulle S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
 RL Nature 387:84-87(1997).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: SOME, TO S.POMBE SPAC23C11.01.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z46728; CAA86704.1; -.
CC PIR; S49790; S49790.
CC GERMOnline; 139625; -.
CC SGD; S0001352; YIL090W.
CC KW Hypothetical protein; Transmembrane.
CC FT TRANSMEM 14 34 POTENTIAL.
CC FT TRANSMEM 41 61 POTENTIAL.
CC FT TRANSMEM 61 95 POTENTIAL.
CC FT TRANSMEM 75 95 POTENTIAL.
CC FT TRANSMEM 150 170 POTENTIAL.
CC FT TRANSMEM 185 205 POTENTIAL.
CC FT TRANSMEM 219 239 POTENTIAL.
CC FT TRANSMEM 295 315 POTENTIAL.
CC FT TRANSMEM 374 394 POTENTIAL.
CC FT TRANSMEM 414 434 POTENTIAL.
CC FT TRANSMEM 462 482 POTENTIAL.
CC SQ SEQUENCE 491 AA; 56332 MW; BAEC65534E04239 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 491;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASSESIS 6
DB 371 ASSESVS 376

RESULT 9
SYT CHLMU
ID - SYT_CHLMU STANDARD; PRT; 635 AA.
AC Q9JF9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine--tRNA ligase)
DE (ThRS).
GN THRS OR TC0870.
OS Chlamydia muridarum.
CC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni99;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Petersen J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-threonine + tRNA(Thr) = AMP +
CC diphosphate + L-threonyl-tRNA(Thr)
CC -1- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AE002353; AAP39666.1; -.
CC PIR; A81656; A81656.
CC HSSP; P00955; LEVL.
CC TIGR; TC0870; -.
CC HAMAP; MF_00184; -. 1.
```

```
DR InterPro; IPR004154; HGTP_anticondon.
DR InterPro; IPR004095; TGS_dom.
DR InterPro; IPR002314; tRNA-synt_2b.
DR InterPro; IPR002320; tRNA-synt_thr.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF03129; HGTP_anticondon; 1.
DR Pfam; PF02824; TGS; 1.
DR Pfam; PF00587; tRNA-synt_2b; 1.
DR PRINTS; PRO1647; TRNASYNTHTR.
DR TIGRFAMs; TIGR00418; thrs; 1.
DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
FT DOMAIN 237 528 CATALYTIC.
FT METAL 328 328 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 379 379 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 505 505 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 635 AA; 72554 MW; 39825894E70521AC CRC64;

Query Match 96.0%; Score 24; DB 1; Length 635;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASSESIS 6
DB 568 ASSESVS 573

RESULT 10
BOSS_DROVI
ID - BOSS_DROVI STANDARD; PRT; 893 AA.
AC Q24738;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Bride of sevenless protein precursor.
GN BOSS.
OS Drosophila virilis (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281693; PubMed=8506350;
RA Hart A.C., Harrison S.D., van Vactor D.L. Jr., Rubin G.M.,
RA Zipursky S.L.;
RT "The interaction of bride of sevenless with sevenless is conserved
RT between Drosophila virilis and Drosophila melanogaster."
RL Proc. Natl. Acad. Sci. U.S.A. 90:5047-5051(1993).
CC -1- FUNCTION: Acts as a ligand for sevenless tyrosine-kinase receptor
CC during eye development.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L08132; AAA72332.1; -.
CC PIR; A47550; A47550.
CC FlyBase; FBgn0013105; Dvir\boss.
DR InterPro; IPR002956; Bride_of_7less.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PRO1223; BRIDE0F7LESS.
KW Transmembrane; Glycoprotein; Vision; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 893 BRIDE OF SEVENLESS PROTEIN.
FT DOMAIN 32 530 EXTRACELLULAR (POTENTIAL).
```

```

FT TRANSMEM 531 551 POTENTIAL.
FT TRANSMEM 563 583 POTENTIAL.
FT TRANSMEM 607 627 POTENTIAL.
FT TRANSMEM 630 650 POTENTIAL.
FT TRANSMEM 653 673 POTENTIAL.
FT TRANSMEM 692 712 POTENTIAL.
FT TRANSMEM 722 742 POTENTIAL.
FT TRANSMEM 752 772 POTENTIAL.
FT DOMAIN 773 893 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 893 AA; 99939 MW; ABB9D4C21DB680C3 CRC64;

Query Match 96.0%; Score 24; DB 1; Length 893;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 676 ASESVS 681

RESULT 11
IG A NEIGO STANDARD; PRT; 1532 AA.
AC P09780;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE IGA-specific serine endopeptidase precursor (EC 3.4.21.72) (IGA
DE protease).
GN IGA.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=MS11;
RX MEDLINE=87115823; PubMed=3027577;
RA Pohlner J., Halter R., Beyreuther K., Meyer T.P.;
RT "Gene structure and extracellular secretion of Neisseria gonorrhoeae
RT IGA protease."
RL Nature 325:458-462(1987).
RN [2]
RP ACTIVE SITE.
RX MEDLINE=90154052; PubMed=2105953;
RA Bachovchin W.W., Plaut A.G., Fienke G.R., Lynch M., Kettner C.A.;
RT "Inhibition of Iga1 proteinases from Neisseria gonorrhoeae and
RT Hemophilus influenzae by peptide prolyl boronic acids."
RL J. Biol. Chem. 265:3738-3743(1990).
CC -!- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.
CC -!- CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC certain Pro-Xaa bonds in the hinge region. No small molecule
CC substrates are known.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS.
CC -!- SIMILARITY: Belongs to peptidase family S6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; X04835; CAA28538.1; -.
DR PIR; A26039; A26039.
DR MEROPS; S06.001; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000710; Peptidase S6.
DR InterPro; IPR004899; Pertactin.
DR InterPro; IPR000824; TrpBP.
DR Pfam; PF03737; Autotransporter; 1.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03212; Pertactin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR TIGRfams; TIGR01414; autotrans_barl; 1.
DR Hydrolase; Serine protease; Zymogen; Autocatalytic cleavage;
KW Transmembrane; Signal.
FT SIGNAL 1 27
FT CHAIN 28 986 IGA-SPECIFIC SERINE ENDOPEPTIDASE.
FT PROPEP 987 1532 HELPER PEPTIDE.
FT ACT SITE 278 278 POTENTIAL.
FT SITE 986 987 CLEAVAGE (AUTO-).
FT SITE 1018 1019 CLEAVAGE (AUTO-).
FT SITE 1121 1122 CLEAVAGE (AUTO-).
SQ SEQUENCE 1532 AA; 168976 MW; 68FF4112BD22F40D CRC64;

Query Match 96.0%; Score 24; DB 1; Length 1532;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 1200 ASESVS 1205

RESULT 12
Y007 CHLTR STANDARD; PRT; 316 AA.
AC O84010;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein CT007.
GN CT007.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -!- SIMILARITY: Belongs to the chlamydial CPN0441/CT007/TC0275 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001275; AAC67597.1; -.
DR PIR; B71569; B71569.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 316 AA; 35637 MW; 904B2CC28F5CB06E CRC64;

Query Match 92.0%; Score 23; DB 1; Length 316;
Best Local Similarity 83.3%; Pred. No. 73;

```

```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASESIS 6
DB 172 ASESLS 177

RESULT 13
MGEL_HUMAN STANDARD; PRT; 373 AA.
ID MGEL_HUMAN
AC Q9UBF1; Q9PLM6; Q9PLM7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Melanoma-associated antigen E1 (MAGE-E1 antigen) (MAGE-C2 antigen)
DE (Hepatocellular carcinoma-associated antigen 587) (Cancer-testis
DE antigen CT10).
GN MAGEE1 OR MAGEC2 OR HCA587.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20164600; PubMed=10699956;
RA Gure A.O., Stockert E., Arden K.C., Boyer A.D., Viars C.S.,
RA Scanlan M.J., Old L.J., Chen Y.-T.;
RT "CT10: a new cancer-testis (CT) antigen homologous to CT7 and the MAGE
RT family, identified by representational difference analysis.";
RL Int. J. Cancer 85:726-732(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20321428; PubMed=10861452;
RA Lucas S., De Plaen E., Boon T.;
RT "MAGE-B5, MAGE-B6, MAGE-C2, and MAGE-C3: four new members of the MAGE
RT family with tumor-specific expression.";
RL Int. J. Cancer 87:55-60(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Hepatoma;
RA Wang Y., Han X.-J., Pang X.-W., Vaughan H.A., Qu W., Dong X.-Y.,
RA Peng J.-R., Zhao H.-T., Rui J.-A., Leng X.-S., Cebon J., Burgess A.W.,
RA Chen W.-F.;
RT "Large scale identification of human hepatocellular carcinoma-
RT associated antigens by autoantibodies.";
RL J. Immunol. 169:1102-1109(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Skin;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Expressed in testis and in liver cancers.
CC -1- SIMILARITY: Contains 1 MAGE domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF116194; AAF34816.1; -
DR EMBL; AF116195; AAF34817.1; -
DR EMBL; AF196483; AAF07211.1; -
DR EMBL; AF196482; AAF07210.1; -
DR EMBL; AF151378; AAF36533.1; -
DR EMBL; AF239802; AAK15073.1; -
DR EMBL; BC005891; AAH05891.1; -
DR Genew; HGNC:13574; MAGEE1.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS00838; MAGE; 1.
KW Antigen; Multigene family.
FT DOMAIN 141 336 MAGE.
FT DOMAIN 34 38 POLY-GLU.
FT DOMAIN 40 60 SER-RICH.
FT DOMAIN 67 70 POLY-GLU.
FT DOMAIN 82 118 SER-RICH.
FT CONFLICT 189 189 M -> T (IN REF. 1; AAF34816).
FT CONFLICT 334 334 K -> E (IN REF. 1).
SQ SEQUENCE 373 AA; 41162 MW; D55867D7972BF49F CRC64;

Query Match 92.0%; Score 23; DB 1; Length 373;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
DB 358 ASESLS 363

RESULT 14
LMRA_LACLA STANDARD; PRT; 590 AA.
AC Q9CHL8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Multidrug resistance ABC transporter ATP-binding and permease
DE protein.
GN LMRA OR LI0711.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LI1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Winkler P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis LI1403.";
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: Efflux transporter for a variety of amphiphilic cationic
CC compounds, including antibiotics (By similarity).
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB006305; AAK04809.1; ALT_INIT.
DR FIR; G86713; G86713.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00929; ABC_TMIF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Transport; Transmembrane; Antibiotic resistance;
KW Complete proteome.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 292 312 POTENTIAL.
FT DOMAIN 349 580 ABC_TRANSPORTER.
FT NP_BIND 382 389 ATP (POTENTIAL).
SQ SEQUENCE 590 AA; 64698 MW; 8DF41419A01388CA CRC64;

Query Match 92.0%; Score 23; DB 1; Length 590;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 214 ASESLS 219

RESULT 15
LMRA_LACLC STANDARD; PRT; 590 AA.
AC P97046;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Multidrug resistance ABC transporter ATP-binding and permease
DE protein.
GN LMRA.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]_TaxID=1359;
RP SEQUENCE FROM N.A.
RC STRAIN=MG1363;
RX MEDLINE=97008061; PubMed=8855237;
RA van Veen H.W., Venema K., Bolhuis H., Oussenko I., Kok J., Poolman B.,
RT Driessen A.J., Konings W.N.;
RT "Multidrug resistance mediated by a bacterial homolog of the human
RT multidrug transporter MDR1.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10668-10672 (1996).
CC -!- FUNCTION: Efflux transporter for a variety of amphiphilic cationic
CC compounds, including antibiotics.
CC -!- SUBUNIT: Homodimer (potential).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (probable).
CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U63741; AAB49750.1; -.
DR InterPro; IPR003593; AAA_ATPase.
```

```
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00929; ABC_TMIF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Transport; Transmembrane; Antibiotic resistance.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 292 312 POTENTIAL.
FT DOMAIN 349 580 ABC_TRANSPORTER.
FT NP_BIND 382 389 ATP (POTENTIAL).
SQ SEQUENCE 590 AA; 64617 MW; CCE4E7D4EF912B8D CRC64;

Query Match 92.0%; Score 23; DB 1; Length 590;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 214 ASESLS 219

Search completed: October 6, 2004, 16:30:30
Job time : 6.68421 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23:13 ; Search time 20.1053 Seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-10
Perfect score: 25
Sequence: 1 ASESIS 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 25 | 100.0 | 225 | 16 | Q98CQ2 rhizobium 1 |
| 2 | 25 | 100.0 | 248 | 16 | Q98CQ6 mycoplasma |
| 3 | 25 | 100.0 | 266 | 11 | Q98M54 mus musculus |
| 4 | 25 | 100.0 | 266 | 11 | Q88837 mus musculus |
| 5 | 25 | 100.0 | 298 | 10 | Q81656 hemerocalli |
| 6 | 25 | 100.0 | 362 | 16 | Q88K39 pseudomonas |
| 7 | 25 | 100.0 | 425 | 10 | Q7X6H5 oryza sativ |
| 8 | 25 | 100.0 | 449 | 4 | Q86WU4 homo sapien |
| 9 | 25 | 100.0 | 483 | 5 | Q86FM3 caenorhabdi |
| 10 | 25 | 100.0 | 491 | 16 | Q828Y7 streptomyce |
| 11 | 25 | 100.0 | 543 | 2 | Q83ZD8 staphylococ |
| 12 | 25 | 100.0 | 565 | 17 | Q9HJX3 thermoplasm |
| 13 | 25 | 100.0 | 602 | 16 | Q7VDF7 prochloroco |
| 14 | 25 | 100.0 | 602 | 16 | Q7V2Q1 prochloroco |
| 15 | 25 | 100.0 | 604 | 16 | Q7V8S4 prochloroco |
| 16 | 25 | 100.0 | 649 | 3 | Q94538 schizosacch |

| | | | | | | |
|----|----|-------|------|----|--------|---------------------|
| 17 | 25 | 100.0 | 732 | 3 | O94462 | O94462 schizosacch |
| 18 | 25 | 100.0 | 797 | 4 | Q9UQM5 | Q9UQM5 homo sapien |
| 19 | 25 | 100.0 | 820 | 11 | Q8CI04 | Q8CI04 mus musculu |
| 20 | 25 | 100.0 | 840 | 11 | Q8COS5 | Q8COS5 mus musculu |
| 21 | 25 | 100.0 | 970 | 16 | Q8DYL7 | Q8DYL7 streptococ |
| 22 | 25 | 100.0 | 999 | 16 | Q8P3V6 | Q8P3V6 xanthomonas |
| 23 | 25 | 100.0 | 1004 | 16 | Q8PFD5 | Q8PFD5 xanthomonas |
| 24 | 25 | 100.0 | 1179 | 4 | Q13545 | Q13545 homo sapien |
| 25 | 25 | 100.0 | 1188 | 4 | Q92656 | Q92656 homo sapien |
| 26 | 25 | 100.0 | 1188 | 4 | Q92835 | Q92835 homo sapien |
| 27 | 25 | 100.0 | 1189 | 4 | Q9UE80 | Q9UE80 homo sapien |
| 28 | 25 | 100.0 | 1189 | 4 | Q00145 | Q00145 homo sapien |
| 29 | 25 | 100.0 | 1310 | 16 | Q8E473 | Q8E473 streptococ |
| 30 | 25 | 100.0 | 1320 | 11 | Q9JK25 | Q9JK25 rattus norv |
| 31 | 25 | 100.0 | 1349 | 4 | Q8WW11 | Q8WW11 homo sapien |
| 32 | 25 | 100.0 | 1391 | 11 | Q922J3 | Q922J3 mus musculu |
| 33 | 25 | 100.0 | 4318 | 5 | Q27802 | Q27802 triphneustes |
| 34 | 24 | 96.0 | 142 | 16 | Q8ESH1 | Q8ESH1 streptococ |
| 35 | 24 | 96.0 | 168 | 16 | Q8A2V6 | Q8A2V6 bacteroides |
| 36 | 24 | 96.0 | 175 | 2 | Q524B3 | Q524B3 escherichia |
| 37 | 24 | 96.0 | 180 | 16 | Q8NR50 | Q8NR50 corynebacte |
| 38 | 24 | 96.0 | 304 | 17 | Q9HST0 | Q9HST0 halobacteri |
| 39 | 24 | 96.0 | 349 | 5 | Q18213 | Q18213 caenorhabdi |
| 40 | 24 | 96.0 | 352 | 16 | Q894N9 | Q894N9 clostridium |
| 41 | 24 | 96.0 | 398 | 5 | Q8T909 | Q8T909 drosophila |
| 42 | 24 | 96.0 | 401 | 5 | Q8IPW9 | Q8IPW9 drosophila |
| 43 | 24 | 96.0 | 440 | 5 | Q9VQL5 | Q9VQL5 drosophila |
| 44 | 24 | 96.0 | 441 | 5 | Q9NKM1 | Q9NKM1 dictyosteli |
| 45 | 24 | 96.0 | 443 | 9 | O21860 | O21860 streptococ |

ALIGNMENTS

RESULT 1

Q98CQ2 PRELIMINARY; PRT; 225 AA.
AC Q98CQ2:
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein ml15052.
GN ML15052.

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAFF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kato T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,

RA Takeuchi C., Yanada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti.";

RL DNA Res. 7:331-338(2000).

DR EMBL; AP003005; BAB51569.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 225 AA; 25368 MW; 9149D55CEB83A3877 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 225;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6

DB 154 ASESIS 159

RESULT 2

Q98QK6
ID Q98QK6 PRELIMINARY; PRT; 248 AA.
AC Q98QK6
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein MYPJ_3250.
GN MYPJ_3250.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445564; CAC13498.1; --
DR PIR; E90552; E90552.
DR Mypulist; MYPJ_3250; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 248 AA; 2855 MW; D74C9AA260DA049C CRC64;

Query Match 100.0%; Score 25; DB 16; Length 248;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 11 ASESIS 16

RESULT 3
Q99M54
ID Q99M54 PRELIMINARY; PRT; 266 AA.
AC Q99M54
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Gene rich cluster, C8 gene.
GN GRCC8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002006; AAH02006.1; --
DR MGD; MGI:1315198; Grcc8.
SQ SEQUENCE 266 AA; 28573 MW; 6E0845BA96CF3C57 CRC64;

Query Match 100.0%; Score 25; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 100 ASESIS 105

RESULT 4
O88837
ID O88837 PRELIMINARY; PRT; 266 AA.
AC O88837
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

Chromosome 6 BAC-284H12 (RESEARCH GENETICS mouse BAC LIBRARY) complete
sequence (RESEARCH GENETICS mouse BAC LIBRARY) (241005A12RIK protein)
(Gene rich cluster).
GN GRCC8 OR C8 OR 241005A12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
Lu J., Gorrell J.H., Chinault A.C., Belmont J.W., Miller W.,
RA Gibbs R.A.;
RT "Comparative sequence analysis of a gene-rich cluster at human
RT chromosome 12p13 and its syntenic region in mouse chromosome 6.";
RL Genome Res. 8:29-40(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Muzny D., Ansari-Lari M.A., Timms K.M., Yu W., Dugan S., Lu J.,
RA Shen Y., Rowland K., Liu W., Perez L., Ding Y., Haywood M., Jain A.,
RA Leal B., Logan O., Nguyen V., Savage L., Shen H., Worley K., Chen E.,
RA Forcum J., Arenson A.D., Chiu M.W., Gorrell J.H., Brundage E., Di W.,
RA Chinault C., Nelson D., Gibbs R.A.;
RT "Direct Submission.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Muzny D., Arenson A.D., Adams C., Brundage E., Bunac C., Carvelli K.,
RA Chacko J., Chen J., Di W., Ding Y., Dugan S., Durbin J., Forcum J.,
RA Ganesh R., Garcia C., Goodman M., Gorrell J.H., Haywood M.,
RA Hernandez J., Jackson L., Jin S., Kampal R., Karpathy S., Kovar C.,
RA Leal B., Li Y., Lichtarge O., Liu W., Logan O., Lu J., Ly T.,
RA Martinez C., Oswal G., Perez L., Rashid N.D., Rowland K., Savage L.,
RA Scherer S.E., Shen H., Simon M., Stovall K., Timms K.M., Todd J.,
RA Vo Q., Williamson A., Worley K.C., Yu W., Chinault C., Nelson D.,
RA Gibbs R.A.;
RT "Direct Submission.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX TISSUE=Embryo, Embryonic stem cells, and small intestine;
MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

```

DR EMBL; AC002397; AAC36014.1; -
DR EMBL; AK011313; BAB27539.1; -
DR EMBL; AK008606; BAB25773.1; -
DR EMBL; AK010409; BAB26616.1; -
DR EMBL; AK028159; BAB25785.1; -
DR MG2; MGI:1315198; Grc8.
SQ SEQUENCE 266 AA; 28708 MW; 8E02E77ACAAAC2F8 CRC64;

Query Match 100.0%; Score 25; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 100 ASESIS 105

RESULT 5
O81656 PRELIMINARY; PRT; 298 AA.
AC O81656;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Senescence-associated protein 6.
GN SAG.
OS Hemerocallis sp. (Daylily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC Hemerocallidaceae; Hemerocallis.
ON NCBI_TaxID=29711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Stella d'Oro; TISSUE=Senescing petals;
RX MEDLINE=9939248; PubMed=10412903;
RA Panavas T., Pikula A., Reid P.D., Rubinstein B., Walker E.L.;
RT "Identification of senescence-associated genes from daylily petals.";
RL Plant Mol. Biol. 40:237-248(1995).
DR EMBL; AF082031; AAC34856.1; -
DR HSSP; P24289; IAK0.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006308; P:DNA catabolism; IEA.
DR InterPro; IPR008947; PLC Nuclease.
DR InterPro; IPR003154; S1/Finuclease.
DR Pfam; PF02955; Nuclease; 1.
DR SEQUENCE 298 AA; 34109 MW; 7FF90F476FFB0057 CRC64;

Query Match 100.0%; Score 25; DB 10; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 235 ASESIS 240

RESULT 6
Q88K39 PRELIMINARY; PRT; 362 AA.
AC Q88K39;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE L-asparaginase II.
GN ANSA OR PP2453.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
ON NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzaz A., Hance I.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoef A., Tuemmler B.,
RA Fraser C.M.;
RA "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016783; AAN68085.1; -
DR TIGR; PF2453; -
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR006034; Asp/Glutamase.
DR Pfam; PF00710; Asparaginase; 1.
DR PRINTS; PS00139; ASNGLNASE.
DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 362 AA; 38608 MW; C33F185DB17053F0 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 91 ASESIS 96

RESULT 7
Q7X6H5 PRELIMINARY; PRT; 425 AA.
ID Q7X6H5;
AC Q7X6H5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNBA0027P08.5 protein (OSUNBA0058G03.8 protein).
GN OSUNBA0027P08.5 OR OSUNBA0058G03.8.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
ON NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731593; CAD40973.1; -
DR EMBL; AL731606; CAD41047.1; -
SQ SEQUENCE 425 AA; 47596 MW; CE77D377938AAAF CRC64;

Query Match 100.0%; Score 25; DB 10; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 25 ASESIS 30

RESULT 8
Q86WU4

```

ID Q86WU4 PRELIMINARY; PRT; 449 AA.
 AC Q86WU4;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC047895; AAH47895.1; -.
 DR InterPro; IPR000938; CAP_GLY.
 DR Pfam; PF01302; CAP_GLY; 2.
 DR PROSITE; PS00845; CAP_GLY_1; 2.
 DR PROSITE; PS02445; CAP_GLY_2; 2.
 KW Hypothetical protein.
 FT NON_TER 449 449
 SQ SEQUENCE 449 AA; 48881 MW; 2C46D1443E1438A8 CRC64;

Query Match 100.0%; Score 25; DB 4; Length 449;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 1 ASESIS 6
 Db 192 ASESIS 197

RESULT 9

Q86FM3 PRELIMINARY; PRT; 483 AA.
 AC Q86FM3;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein Y37ELLAR.7.
 GN Y37ELLAR.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium. ";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Miller N., Maggi L.;
 RT "The sequence of C. elegans cosmid Y37ELLAR.7";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;

RA Waterston R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024759; AAP13762.1; -.
 DR WormPep; Y37ELLAR.7; CE33855.
 DR InterPro; IPR007110; Ig-like.
 KW Hypothetical protein.
 SQ SEQUENCE 483 AA; 55679 MW; AB174C4F622E3P46 CRC64;

Query Match 100.0%; Score 25; DB 5; Length 483;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 1 ASESIS 6
 Db 138 ASESIS 143

RESULT 10

Q828Y7 PRELIMINARY; PRT; 491 AA.
 AC Q828Y7;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Putative GTP-binding protein.
 GN SAV6524.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites. ";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=1262562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis. ";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005047; BAC74235.1; -.
 DR GO; GO:000166; F:nucleotide binding; IEA.
 DR InterPro; IPR003593; AAA_ATase.

```

DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PRO0326; GTP1_OBG.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00650; MG442; 2.
DR TIGRFAMs; TIGR00231; small_GTP; 2.
KW Complete proteome.
SQ SEQUENCE 491 AA; 5353 MW; 7BAD8E0449D14AC CRC64;

Query Match 100.0%; Score 25; DB 16; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 319 ASESIS 324

RESULT 11
Q83ZD8 Q83ZD8 PRELIMINARY; PRT; 543 AA.
AC Q83ZD8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Hypothetical protein.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JGSC 4469;
RA Ma X., Ito T., Okuma K., Hiramatsu K.;
RT "Historical distribution of SCCmec allotype in healthcare-associated
RT MRSA strains in Japan and France.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB097677; BAC76058.1; -.
KW Hypothetical protein.
SQ SEQUENCE 543 AA; 62077 MW; 9CD7B3183EBF075 CRC64;

Query Match 100.0%; Score 25; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 413 ASESIS 418

RESULT 12
Q9HJX3 Q9HJX3 PRELIMINARY; PRT; 565 AA.
AC Q9HJX3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical membrane protein.
GN TA0839.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatiales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";

RL Nature 407:508-513 (2000).
DR EMBL; AL445065; CAC11968.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008915; Peptidase_M50.
DR Pfam; PF02163; Peptidase_M50; 1.
KW Complete proteome.
SQ SEQUENCE 565 AA; 61966 MW; B78E9738C56A9B01 CRC64;

Query Match 100.0%; Score 25; DB 17; Length 565;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 440 ASESIS 445

RESULT 13
Q7VDF7 Q7VDF7 PRELIMINARY; PRT; 602 AA.
AC Q7VDF7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Membrane GTPase LepA.
GN LEPA OR PRO0419.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=12119;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / CCMP 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486;
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.I., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxypototrophic genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025 (2003).
DR EMBL; AB017162; AAP99485.1; -.
KW Complete proteome.
SQ SEQUENCE 602 AA; 67256 MW; 659E639EBF27D290 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 602;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
Db 544 ASESIS 549

RESULT 14
Q7VZQ1 Q7VZQ1 PRELIMINARY; PRT; 602 AA.
AC Q7VZQ1
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GTP-binding protein LepA.
GN LEPA OR FMM0420.
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22825698; PubMed=12917642;
RX Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,

```

RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic
 RT niche differentiation.";
 RL Nature 424:1042-1047(2003).
 DR EMBL; BX572091; CAE18879.1; --
 KW Complete proteome.
 SQ SEQUENCE 604 AA; 67457 MW; 5916A8E964658E90 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 602;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 Db 544 ASESIS 549

RESULT 15

QV78S4
 ID QV78S4 PRELIMINARY; PRT; 604 AA.
 AC QV78S4;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE GTP-binding protein LepA.
 GN LEPA OR PNT0257.
 OS *Prochlorococcus marinus* (strain MIT 9313).
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC *Prochlorococcus*.
 OX NCBI_TaxID=74547;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22825698; PubMed=12917642;
 RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
 RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic
 RT niche differentiation.";
 RL Nature 424:1042-1047(2003).
 DR EMBL; BX572095; CAE20432.1; --
 KW Complete proteome.
 SQ SEQUENCE 604 AA; 67299 MW; F1D561845D76A804 CRC64;

Query Match 100.0%; Score 25; DB 16; Length 604;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASESIS 6
 Db 544 ASESIS 549

Search completed: October 6, 2004, 16:34:02
 Job time : 25.1053 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:23.13 ; Search time 30.1579 Seconds
(without alignments)
94.160 Million cell updates/sec

Title: US-09-635-974A-12

Perfect score: 56

Sequence: 1 QQNNNWPTT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 40 | 71.4 | 160 | Q8RP67 | Q8RP67 anaplasma p |
| 2 | 40 | 71.4 | 184 | Q8R011 | Q8R011 mus musculus |
| 3 | 40 | 71.4 | 421 | Q7TP45 | Q7TP45 rattus norv |
| 4 | 40 | 71.4 | 972 | Q9AAL8 | Q9AAL8 caulobacter |
| 5 | 38 | 67.9 | 598 | Q8ZGM7 | Q8ZGM7 yersinia pe |
| 6 | 38 | 67.9 | 1109 | Q9VTB8 | Q9VTB8 drosophila |
| 7 | 37 | 66.1 | 111 | Q8V3P6 | Q8V3P6 swinepox vi |
| 8 | 37 | 66.1 | 130 | Q8KQ00 | Q8KQ00 anaplasma p |
| 9 | 37 | 66.1 | 134 | Q83VV5 | Q83VV5 anaplasma p |
| 10 | 37 | 66.1 | 139 | Q83VU7 | Q83VU7 anaplasma p |
| 11 | 37 | 66.1 | 153 | Q8KR43 | Q8KR43 vibrio chol |
| 12 | 37 | 66.1 | 166 | Q84IQ6 | Q84IQ6 anaplasma p |
| 13 | 37 | 66.1 | 178 | Q8RPC0 | Q8RPC0 anaplasma p |
| 14 | 37 | 66.1 | 452 | Q8EX49 | Q8EX49 mycoplasma |
| 15 | 37 | 66.1 | 503 | Q8F9N5 | Q8F9N5 leptospira |
| 16 | 37 | 66.1 | 866 | Q9VMY4 | Q9VMY4 human immun |

ALIGNMENTS

RESULT 1

Q8RP67 ID Q8RP67 PRELIMINARY; PRT; 160 AA.

AC Q8RP67;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Major surface protein hypervariable region (Fragment).
GN MSP2.

OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OX Anaplasmataceae; Anaplasma.
OC NCBI_taxID=948;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Webster;

RX MEDLINE=21843110; PubMed=11854205;

RA Caspersen K., Park J.H., Patil S., Dumler J.S.;

RT "Genetic variability and stability of Anaplasma phagocytophila msp2

RL Infect. Immun. 70:1230-1234 (2002).

DR EMBL; AF443406; AAJ91714.1; -

DR InterPro; IPR002566; Surface_Ag_msp4.

DR Pfam; PF01617; Surface_Ag_2; 1.

FT NON_TER 1 160

FT NON_TER 1 160

SQ SEQUENCE 160 AA; 16602 MW; 001D7481ECA85AC CRC64;

Query Match 71.4%; Score 40; DB 2; Length 160;

Best Local Similarity 75.0%; Pred. No. 30;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQNNNWPTT 8

Db 85 QDNKNWPT 92

RESULT 2

Q8R011

Q84FL0 pantoa agg
Q8T8M4 caenorhabdi
Q9P3N7 neurospora
P91350 caenorhabdi
Q26812 methanobact
Q91B29 citrus yell
Q8Xym3 trichomonas
Q9u183 homo sapien
Q83vv6 anaplasma p
Q8kqp3 anaplasma p
Q8kqg2 anaplasma p
Q8ei1 shewanella
Q8kq50 anaplasma p
Q84lp8 anaplasma p
Q8rp60 anaplasma p
Q8rp63 anaplasma p
Q8rpc2 anaplasma p
Q8rpc4 anaplasma p
Q8rbp9 anaplasma p
Q87tp0 clostridium
Q81218 caenorhabdi
O16101 drosophila
Q9vmx7 drosophila
Q8szg4 drosophila
Q8rbp4 anaplasma p
Q84967 anaplasma p
Q7vni2 haemophilus
Q88spi pseudomonas
Q814G9 caenorhabdi

17 37 66.1 877 2 Q84FL0
18 37 66.1 889 5 Q8T8M4
19 37 66.1 1140 3 Q9P3N7
20 37 66.1 1562 5 P91350
21 37 66.1 1755 17 Q26812
22 37 66.1 1983 12 Q91B29
23 37 66.1 2320 5 Q8Xym3
24 36 64.3 108 4 Q9u183
25 36 64.3 130 2 Q83vv6
26 36 64.3 135 2 Q8kqp3
27 36 64.3 142 2 Q8kqg2
28 36 64.3 154 16 Q8ei1
29 36 64.3 167 2 Q8kq50
30 36 64.3 167 2 Q84lp8
31 36 64.3 170 2 Q8rp60
32 36 64.3 170 2 Q8rp63
33 36 64.3 174 2 Q8rpc2
34 36 64.3 179 2 Q8rpc4
35 36 64.3 182 2 Q8rbp9
36 36 64.3 205 16 Q87tp0
37 36 64.3 265 5 O16101
38 36 64.3 265 5 Q9vmx7
39 36 64.3 272 5 Q8szg4
40 36 64.3 323 2* Q8rbp4
41 36 64.3 326 2 Q84967
42 36 64.3 360 16 Q7vni2
43 36 64.3 397 16 Q88spi
44 36 64.3 399 5 Q814G9
45 36 64.3


```

GN CC0579
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C.; Feldblyum T.V.; Laub M.T.; Paulsen I.T.; Nelson K.E.;
RA Eisen J.; Heidelberg J.F.; Alley M.R.K.; Ohta N.; Maddock J.R.;
RA Potocka I.; Nelson W.C.; Newton A.; Stephens C.; Phadke N.D.; Ely B.;
RA DeBoy R.T.; Dodson R.J.; Durkin A.S.; Gwinn M.L.; Haft D.H.;
RA Konovalov J.F.; Smit J.; Craven M.B.; Khouri H.; Shetty J.; Berry K.;
RA Uterback T.; Tran K.; Wolf A.; Vamathevan J.; Ermolaeva M.; White O.;
RA Salzberg S.L.; Venter J.C.; Shapiro L.; Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
RL EMBL; AE005732; AA022565.1; -.
DR PIR; A87321; A87321.
DR TIGR; CC0579; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00593; TonB_dep_Reg; 1.
KW Receptor; Complete Proteome.
SQ SEQUENCE 972 AA; 104732 MW; B599CDD10C26AC4D CRC64;

Query Match 71.4%; Score 40; DB 16; Length 972;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNNNWPTT 9
DB 390 ENPNWPTT 397
| | | | |
| | | | |

RESULT 5
O82GM7 PRELIMINARY; PRT; 598 AA.
AC O82GM7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Putative membrane protein (Putative sulfatase).
GN YP01260 OR Y2924.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J.; Wren B.W.; Thomson N.R.; Tibball R.W.; Holden M.T.G.;
RA Prentice M.B.; Sebahia M.; James K.D.; Churcher C.; Mungall K.L.;
RA Baker S.; Basham D.; Bentley S.D.; Brooks K.; Cerdeno-Tarraga A.M.;
RA Chillingworth T.; Cronin A.; Davies R.M.; Davis P.; Dougan G.;
RA Feltwell T.; Hamlin N.; Holroyd S.; Jagels K.; Karlyshev A.V.;
RA Leather S.; Moule S.; Oyston P.C.F.; Quail M.; Rutherford K.;
RA Simmonds M.; Skelton J.; Stevens K.; Whitehead S.; Barrall B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W.; Burland V.; Plunkett G. III; Boutin A.; Mayhew G.F.; Liss P.;
RA Ferra N.T.; Rose D.J.; Mau B.; Zhou S.; Schwartz D.C.;
RA Fetherston J.D.; Lindler L.E.; Brubaker R.R.; Plano G.V.;
RA Straley S.C.; McDonough K.A.; Nilles M.L.; Matson J.S.; Blattner F.R.;

```

```

RA Perry R.D.;
RT "Genome sequence of Versinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AJ414147; CAC90094.1; -.
DR EMBL; AE013894; AAM86475.1; -.
DR PIR; AC0154; AC0154.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0008484; F:sulfuric ester hydrolase activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:metabolism; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000917; Sulfatase.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 598 AA; 67474 MW; C6EE2A0BDC1E43A2 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 598;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9
DB 534 QRNNWVAT 542
| | | | |
| | | | |

RESULT 6
Q9VTB8
ID Q9VTB8 PRELIMINARY; PRT; 1109 AA.
AC Q9VTB8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CG7958 protein (LDL692lp).
GN TNA OR CG7958.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA Sutton G.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Sutton R.C.; Rogers Y.-H.C.; Blazer R.G.; Zhang Q.; Chen L.X.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;
RA Abriel J.F.; Agbayani A.; An H.-J.; Andrews-Pfannkoch C.; Baldwin D.;
RA Balwle R.M.; Basu A.; Baxendale J.; Bayraktoglu L.; Beasley E.M.;
RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;
RA Borkova D.; Botchan M.R.; Bouck J.; Brokstein P.; Brotter P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cwayle S.; Dahlke C.; Davenport L.B.; Davies P.;
RA de Pablos B.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.J.; Dou L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferreira S.; Fleischmann W.;
RA Fowler C.; Gabriellian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Glodek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.; Heiman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;
RA Jalali M.; Kalush F.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;
RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;
RA Lasko P.; Lei Y.; Levitsky A.A.; Li J.; Li Z.; Liang Y.; Lin X.;
RA Liu X.; Mattie B.; McIntosh T.C.; McLeod M.P.; McPherson D.;
RA Mestulov G.; Milshina N.V.; Mobarry C.; Morris J.; Moshrefi A.;
RA Mount S.M.; Moy M.; Murphy B.; Murphy L.; Muzny D.M.; Nelson D.L.;
RA Nelson D.R.; Nelson K.A.; Nixon K.; Nusskern D.R.; Pacieb J.M.;
RA Palazzolo M.; Pittman G.S.; Pan S.; Pollard J.; Puri V.; Reese M.G.;

```

RA Reinert K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tredler C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassmann D.A., Weirick G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banson J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Mirza S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Begman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO03547; AAF50134.2; -;
DR EMBL; BT003272; AAO25029.1; -;
DR FlyBase; FBgn036103; tna.
DR GO; GO:0008270; F-actin ion binding; IEA.
DR InterPro; IPR004181; Znf_MIZ.
DR Pfam; PF02891; zf-MIZ; 1.
SQ SEQUENCE 1109 AA; 116491 MW; 4FD726183E642AC CRC64;
Query Match 67.9%; Score 38; DB 5; Length 1109;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 QONNNWPT 9
Db 591 QMNTNWPT 599

RESULT 7
Q8V3P6 PRELIMINARY; PRT; 111 AA.
AC Q8V3P6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SPV049 hypothetical protein.
GN SPV049.
OS Swinepox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Suipoxvirus.
OX NCBI_TaxID=10276;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17077-99;
RX MEDLINE=21624277; PubMed=11752168;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Osorio F.A., Balinsky C.,
RA Kutish G.F., Rock D.L.;
RT "The genome of swinepox virus."
RL J. Virol. 76:783-790(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=17077-99;
RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C., Osorio F.A., Zsak L.,
RA Kutish G.F., Rock D.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410153; AAL69787.1; -;
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0005509; F-actin ion binding; IEA.
DR GO; GO:0007156; Phospholipid cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
KW Hypothetical protein.
SQ SEQUENCE 111 AA; 13043 MW; AF715FFF23E704E8 CRC64;
Query Match 66.1%; Score 37; DB 12; Length 111;
Best Local Similarity 87.5%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 QONNNWPT 8
Db 45 QONNNIPT 52

RESULT 8
Q8V3P6 PRELIMINARY; PRT; 130 AA.
AC Q8V3P6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 44 kDa major outer membrane protein (Fragment).
GN P44-30.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22144240; PubMed=12149362;
RA Lin Q., Zhi N., Ohashi N., Horowitz H.W., Aguerro-Rosenfeld M.E.,
RA Raffalli J., Wormser G.P., Rikihisa Y.;
RT "Analysis of Sequences and Loci of p44 Homologs Expressed by Anaplasma
phagocytophila in Acutely Infected Patients."
RL J. Clin. Microbiol. 40:2981-2988(2002).
DR EMBL; AY064519; AAL78182.1; -;
FT NON TER 1
FT NON TER 130
SQ SEQUENCE 130 AA; 13416 MW; 6F735572B031DFD1 CRC64;
Query Match 66.1%; Score 37; DB 2; Length 130;
Best Local Similarity 71.4%; Pred. No. 77;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNWPT 8
:|||||
Db 81 ENKNWPT 87

RESULT 9

Q83VV5 ID Q83VV5 PRELIMINARY; PRT; 134 AA.
AC Q83VV5; MEDLINE=22541496; PubMed=12654783;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Major surface protein 2 (Fragment)
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=var B;
RX MEDLINE=22541496; PubMed=12654783;
RA Barbet A.F., Meens P.F.M., Belanger M., Bowie M.V., Yi J.,
RA Lundgren A.M., Alleman A.R., Wong S.J., Chu F.K., Munderloh U.G.,
RA Jauron S.D.;
RT "Expression of multiple outer membrane protein sequence variants from
a single genomic locus of Anaplasma phagocytophilum.";
RL Infect. Immun. 71:1706-1718(2003).
DR EMBL; AY164497; AA030109.1; -;
FT NON_TER 1
FT NON_TER 134
SQ SEQUENCE 134 AA; 13573 MW; 67E791D5303C15B1 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 134;
Best Local Similarity 71.4%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNWPT 8
:|||||
Db 83 ENKNWPT 89

RESULT 10

Q83VV7 ID Q83VV7 PRELIMINARY; PRT; 139 AA.
AC Q83VV7; MEDLINE=22541496; PubMed=12654783;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Major surface protein 2 (Fragment)
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=var A;
RX MEDLINE=22541496; PubMed=12654783;
RA Barbet A.F., Meens P.F.M., Belanger M., Bowie M.V., Yi J.,
RA Lundgren A.M., Alleman A.R., Wong S.J., Chu F.K., Munderloh U.G.,
RA Jauron S.D.;
RT "Expression of multiple outer membrane protein sequence variants from
a single genomic locus of Anaplasma phagocytophilum.";
RL Infect. Immun. 71:1706-1718(2003).
DR EMBL; AY164505; AA030117.1; -;
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 14576 MW; 6C96C03CC55E121A CRC64;

Query Match 66.1%; Score 37; DB 2; Length 139;
Best Local Similarity 71.4%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 QNNWPT 9
:|||||
Db 89 QNNWPT 95

RESULT 11

Q3KR43 ID Q3KR43 PRELIMINARY; PRT; 153 AA.
AC Q3KR43; MEDLINE=20406833; PubMed=10952301;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VC1803.
GN VC1803.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA McDaniel L., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004257; AAF94952.1; -;
DR PIR; H82153; H82153.
DR TIGR; VC1803; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 153 AA; 17480 MW; C92B7B307501A722 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 153;
Best Local Similarity 62.5%; Pred. No. 91;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNWPT 9
:|||||
Db 65 EENSWPT 72

RESULT 12

Q84IQ6 ID Q84IQ6 PRELIMINARY; PRT; 166 AA.
AC Q84IQ6; MEDLINE=22170783; PubMed=12183586;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Major outer membrane protein SCI-1 (Fragment).
GN P44.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22170783; PubMed=12183586;
RA Igdo J.W., Wu C., Telford S.R., Fikrig E.;
RT "Differential expression of the p44 gene family in the agent of human
granulocytic ehrlichiosis.";
RL Infect. Immun. 70:5295-5298(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Igdo J.W.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF512671; AAP14021.1; -;

```

FT NON_TER 1
FT NON_TER 166
SQ SEQUENCE 166 AA; 17205 MW; 6997EFAD6D745C34 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 166;
Best Local Similarity 71.4%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNWPT 8
DB 101 ENKWPT 107

RESULT 13
QSRPCO PRELIMINARY; PRT; 178 AA.
AC QSRPCO;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Major outer membrane protein p44-16 (Fragment).
GN P44-16.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HZ;
RX MEDLINE=21843103; PubMed=11854198;
RA Zhi N., Onashi N., Tajima T., Mott J., Stich R.W., Grover D.,
RA Telford S.R. III, Lin Q., Rikihisa Y.;
RT "Transcript heterogeneity of the p44 multigene family in a human
RT granulocytic ehrlichiosis agent transmitted by ticks.";
RL Infect Immun 70:1175-1184(2002).
DR EMBL; AF412829; RAL86392.1; -
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
FT NON_TER 178
FT NON_TER 178
SQ SEQUENCE 178 AA; 18129 MW; D05899A564A848F0 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 178;
Best Local Similarity 71.4%; Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNWPT 8
DB 102 ENKWPT 108

RESULT 14
QSRPCO PRELIMINARY; PRT; 452 AA.
AC QSRPCO;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Chromosomal replication initiator protein.
GN MYPE10.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans.";
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AF004170; BAC43791.1; -

```

```

DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0003688; F:DNA replication origin binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006270; P:DNA replication initiation; IEA.
DR GO; GO:0006275; P:regulation of DNA replication; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001957; Bac_DnaA.
DR Pfam; PF00308; bac_dnaA; 1.
DR PRINTS; PR00051; DNaA.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00362; DnaA; 1.
KW Complete proteome.
SQ SEQUENCE 452 AA; 52977 MW; A853FC934F25F0E2 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 452;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QNNWPT 8
DB 137 EENNWKT 144

RESULT 15
QSRPCO PRELIMINARY; PRT; 503 AA.
AC QSRPCO;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA0156.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB011207; AAN47355.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 503 AA; 57820 MW; C50EFFADC2366792 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 503;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QNNWPT 9
DB 280 NNNYPT 286

Search completed: October 6, 2004, 16:34:06
Job time : 34.1579 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:21:13 ; Search time 5.52632 Seconds
(without alignments)
84.800 Million cell updates/sec

Title: US-09-635-974A-12
Perfect score: 56
Sequence: 1 QQNNWPTT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 42 | 75.0 | 108 | 1 KXIM HUMAN | P01605 homo sapien |
| 2 | 42 | 75.0 | 109 | 1 KX3F HUMAN | P01624 homo sapien |
| 3 | 38 | 67.9 | 129 | 1 KX3H HUMAN | P04207 homo sapien |
| 4 | 38 | 67.9 | 320 | 1 CELI AGABI | Q00023 agarius bi |
| 5 | 37 | 66.1 | 275 | 1 KDIU BACSU | P50843 bacillus su |
| 6 | 37 | 66.1 | 1692 | 1 CYAA SCHPO | P14505 schizosacch |
| 7 | 36 | 64.3 | 115 | 1 KV51 MOUSE | P01642 mus musculu |
| 8 | 36 | 64.3 | 3317 | 1 CADN RAT | P58365 ratus norv |
| 9 | 36 | 64.3 | 3354 | 1 CADN MOUSE | Q99P44 mus musculu |
| 10 | 36 | 64.3 | 4910 | 1 KDNI YEAST | Q12019 saccharomyc |
| 11 | 35 | 62.5 | 115 | 1 KX3I HUMAN | P04433 homo sapien |
| 12 | 35 | 62.5 | 224 | 1 RADC XYLFA | Q9P328 xylolla fas |
| 13 | 35 | 62.5 | 224 | 1 RADC XYLFT | Q97f21 xylolla fas |
| 14 | 35 | 62.5 | 354 | 1 PROW ECOLI | P14176 escherichia |
| 15 | 35 | 62.5 | 527 | 1 WALO CHLTR | O84089 chlamydia t |
| 16 | 35 | 62.5 | 591 | 1 FUCI ECOLI | P11552 escherichia |
| 17 | 35 | 62.5 | 1178 | 1 YE56 CABEL | P09747 caenorhabdi |
| 18 | 35 | 62.5 | 2183 | 1 POLG CXB4E | Q86887 c genome po |
| 19 | 35 | 62.5 | 2185 | 1 POLG CXB5P | Q03053 c genome po |
| 20 | 35 | 62.5 | 2194 | 1 POLG EC30B | Q9W078 e genome po |
| 21 | 34 | 60.7 | 102 | 1 CYAV PASMU | P57943 pasteurella |
| 22 | 34 | 60.7 | 112 | 1 PLAS MOUSE | Q9J148 mus musculu |
| 23 | 34 | 60.7 | 241 | 1 COAT CSNV | P14985 chloris str |
| 24 | 34 | 60.7 | 324 | 1 YE64 PASMU | Q9CKY6 pasteurella |
| 25 | 34 | 60.7 | 330 | 1 RIL RAT | P36202 rattus norv |
| 26 | 34 | 60.7 | 366 | 1 MLTA BUCAP | Q8K9A7 buchnera ap |
| 27 | 34 | 60.7 | 452 | 1 GUX1 CRYP | Q00548 cryptonec |
| 28 | 34 | 60.7 | 477 | 1 YBFO ECOLI | P77779 escherichia |
| 29 | 34 | 60.7 | 431 | 1 Y298 MYCGA | P33276 mycoplasma |
| 30 | 34 | 60.7 | 563 | 1 YHJW SALT | P43666 salmonella |
| 31 | 34 | 60.7 | 604 | 1 FUCI HAEIN | P44779 haemophilus |
| 32 | 34 | 60.7 | 725 | 1 YA33 SULSO | Q97297 sulfolobus |
| 33 | 34 | 60.7 | 830 | 1 LEM3 HUMAN | P16109 homo sapien |

| | | | | | |
|----|----|------|------|--------------|---------------------|
| 34 | 34 | 60.7 | 954 | 1 XYNA RUMPL | P29126 ruminococcu |
| 35 | 34 | 60.7 | 1581 | 1 ARO1 PNECA | Q12659 p pentafunc |
| 36 | 34 | 60.7 | 1926 | 1 LPH RABIT | P09849 oryctolagus |
| 37 | 33 | 58.9 | 115 | 1 PLA8 HUMAN | Q9Nzf1 homo sapien |
| 38 | 33 | 58.9 | 150 | 1 MGN ARATH | O23676 arabidopsis |
| 39 | 33 | 58.9 | 231 | 1 NKGC MACMU | Q9PZK6 macaca mula |
| 40 | 33 | 58.9 | 344 | 1 ONPA KLEPN | P24017 klebsiella |
| 41 | 33 | 58.9 | 391 | 1 PER DROIN | P91613 drosophila |
| 42 | 33 | 58.9 | 428 | 1 THRC ECOLI | P00934 escherichia |
| 43 | 33 | 58.9 | 429 | 1 THRC SERMA | P27735 serratia ma |
| 44 | 33 | 58.9 | 457 | 1 NU4M ARBLI | P34941 arabacia lix |
| 45 | 33 | 58.9 | 586 | 1 SYD BORBU | O51402 borrelia bu |

ALIGNMENTS

```
RESULT 1
KVIM HUMAN
ID KVIM HUMAN STANDARD; PRT; 108 AA.
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; PubMed=824717;
RA Capra J.D., Klapper D.G.;
RT "Complete amino acid sequence of the variable domains of two human
RT IGM anti-gamma globulins (Lay/Pom) with shared idiotypic
RT specificities."
RL Scand. J. Immunol. 5:677-684 (1976).
CC -!- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC PIR; A01871; KIHULY.
CC HSSP; P01607; IREI.
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0003823; F:antigen binding; NAS.
CC GO; GO:0006955; P:immune response; NAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;
```

Query Match 75.0%; Score 42; DB 1; Length 108;

Best Local Similarity 77.8%; Pred. No. 1.7;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QQNNWPTT 9

Db 89 QQNNWPTT 97

```

RESULT 2
KV3F_HUMAN
ID KV3F_HUMAN STANDARD; PRT; 109 AA.
AC P01624;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig kappa chain V-III region POM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76276460; PubMed=60899;
RA Klapper D.G., Capra J.D.;
RT "The amino acid sequence of the variable regions of the light chains
RT from two idiotypically cross reactive IgM anti-gamma globulins.";
RL Ann. Immunol. (Paris) 127C:261-271(1976).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
CC PIR; A01897; K3HUPM.
DR HSP; P80362; LWTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region.
FT SIGNAL 1 20 IG KAPPA CHAIN V-III REGION CLL.
FT CHAIN 21 129 FRAMEWORK-1.
FT DOMAIN 21 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 54 FRAMEWORK-2.
FT DOMAIN 55 69 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 70 76 FRAMEWORK-3.
FT DOMAIN 77 108 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 109 118 JKI SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON TER 109 109
SQ SEQUENCE 109 AA; 11922 MW; 62821DDC6A8ABA86 CRC64;

Query Match 75.0%; Score 42; DB 1; Length 109;
Best Local Similarity 77.8%; Pred. No. 1.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQNNWPTT 9
DB 90 QQNNWPTT 98

RESULT 3
KV3H_HUMAN
ID KV3H_HUMAN STANDARD; PRT; 129 AA.
AC P04207;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region CLL precursor (Rheumatoid factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86177570; PubMed=3083417;
RA Jirik F.R., Sarge J., Fong S., Heitzmann J.G., Curd J.G., Chen P.P.,
RA Goldfien R., Carson D.A.;
RT "Cloning and sequence determination of a human rheumatoid factor
RT light-chain gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:2195-2199(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; M12740; AAA58992.1; -.
DR HSP; P80362; LWTL.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION CLL.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 118 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 108 BY SIMILARITY.
FT NON TER 129 129
SQ SEQUENCE 129 AA; 14275 MW; 5C13B411B560CC14 CRC64;

Query Match 57.9%; Score 38; DB 1; Length 129;
Best Local Similarity 85.7%; Pred. No. 9.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQNNWNP 7
DB 109 QQNNWNP 115

RESULT 4
CELL_AGABI
ID CELL_AGABI STANDARD; PRT; 320 AA.
AC Q00023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cellulose-growth-specific protein precursor.
GN CELL.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D649;
RX MEDLINE=93012985; PubMed=1398098;
RA Raguz S., Yaguee E., Wood D.A., Thurston C.F.;
RT "Isolation and characterization of a cellulose-growth-specific gene
RT from Agaricus bisporus.";
RL Gene 119:183-190(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=D649;
RX MEDLINE=94237428; PubMed=8181702;
RA Amesilla A.L., Thurston C.F., Yaguee E.;
RT "CELL: a novel cellulose binding protein secreted by Agaricus
RT bisporus during growth on crystalline cellulose.";
RL FEMS Microbiol. Lett. 116:293-299(1994).
CC -!- FUNCTION: PROBABLE GLYCOSYL HYDROLASE ACTIVE ON CELLULOSE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 fungal-type cellulose-binding (CBD) domain.
CC -!- SIMILARITY: BELONGS TO FAMILY 61 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL; M86356; AAA53434.1; -.
CC PIR; JC1311; JC1311.
CC DR HSP; P00725; 2CBH.
CC DR InterPro; IPR000254; CBD fungal.
CC DR InterPro; IPR005103; Glyco_hydro_61.
CC DR Pfam; PF00734; CBM_1; 1.
CC DR Pfam; PF03443; Glyco_hydro_61; 1.
CC DR SMART; SM00236; fCBP; 1.
CC DR PROSITE; PS00562; CBD FUNGAL; 1.
CC KW Cellulose degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
CC FT SIGNAL 1 29 POTENTIAL.
CC FT CHAIN 30 320 CELLULOSE-GROWTH-SPECIFIC PROTEIN.
CC FT DOMAIN 30 261 CATALYTIC (POTENTIAL).
CC FT DOMAIN 262 285 LINKER (POTENTIAL).
CC FT DOMAIN 286 320 CELLULOSE-BINDING (POTENTIAL).
CC FT DISULFID 292 309 BY SIMILARITY.
CC FT DISULFID 303 319 BY SIMILARITY.
CC FT CARBOHYD 163 163 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SQ SEQUENCE 320 AA; 33754 MW; 5032C8080895CA2B CRC64;

Query Match 67.9%; Score 38; DB 1; Length 320;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNNWPTT 9
DB 162 QNNSWITT 169
|||||

RESULT 5
ID KDUI_BACSU STANDARD; PRT; 275 AA.
AC P50833;

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (BC 5.3.1.17) (5-
DE keto-4-deoxyuronate isomerase) (DKI isomerase).
GN KDUI OR BSU22130.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serron P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kds loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Besterio M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ettian K.D., Errington J., Fabret C., Ferrazi E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly B.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinis S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

```

```

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudsga B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowaka A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viati A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RT Nature 390:249-256(1997).
RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: 4-deoxy-L-threo-5-hexosulose uronate = 3-
CC deoxy-D-glycero-2,5-hexodiolosonate.
CC -!- PATHWAY: Pectin degradation.
CC -!- SIMILARITY: Belongs to the kdul family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL; L47838; AAB38477.1; -.
CC DR EMBL; Z91115; CAB14130.1; -.
CC PIR; E69648; E69648.
CC DR Subtilist; BG11401; kdul.
CC DR HAMAP; MF_00687; -. 1.
CC DR InterPro; IPR007045; KduI.
CC DR Pfam; PF04962; KduI; 1.
CC KW Isomerase; Complete proteome.
CC SQ SEQUENCE 275 AA; 31135 MW; 31B2CE2CE7C54760 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 275;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNNNWPTT 8
DB 183 QNNNWNT 189
|||||

RESULT 6
ID CYAA_SCHPO STANDARD; PRT; 1692 AA.
AC P14605;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
DE cyclase).
GN CYR1 OR SPBC19C7.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90046723; PubMed=2682634;
RA Young D., Riggs M., Field J., Voitek A., Broek D., Wigler M.;
RT "The adenylyl cyclase gene from Schizosaccharomyces pombe.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7989-7993(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345533; PubMed=2668944;
RA Yamawaki-Kataoka Y., Tamaoki T., Choe H.-R., Tanaka H., Kataoka T.;

```

RT "Adenylate cyclases in yeast: a comparison of the genes from
 RL Schizosaccharomyces pombe and Saccharomyces cerevisiae.";
 RN Proc. Natl. Acad. Sci. U.S.A. 86:5693-5697(1989).
 RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wanbut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leilaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA Nature 415:871-880(2002).
 RL The genome sequence of Schizosaccharomyces pombe.";
 RT metabolism by catalyzing the synthesis of a second messenger,
 CC CAMP.
 CC -!- CATALYTIC ACTIVITY: ATP = 3' 5'-cyclic AMP + diphosphate.
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- ENZYME REGULATION: In contrast to yeast cyclase, S.pombe cyclase
 CC is not likely to be regulated by RAS proteins.
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 1 PP2C-like domain.
 CC -!- SIMILARITY: Contains 1 Ras-associating domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M26699; AAA35284.1; -;
 CC EMBL: M24942; AAA35301.1; -;
 CC EMBL: AL023859; CAA19571.1; -;
 CC PIR: A33988; A33988.
 CC GenDB Sponbe; SPBC19C7.03; -;
 CC GO: GO:0007049; P:cell cycle; ISS.
 CC GO: GO:0006139; P:nucleobase, nucleoside, nucleotide and nucl. . .; ISS.
 CC GO: GO:0006163; P:purine nucleotide metabolism; ISS.
 CC GO: GO:0007165; P:signal transduction; ISS.
 CC InterPro: IPR001054; G: cyclase.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR003591; LRR: typ.
 CC InterPro: IPR001932; PP2C-like.
 CC InterPro: IPR000159; RA domain.
 CC Pfam: PF00211; Guanylate_cyc; 1.
 CC Pfam: PF00560; LRR; 11.
 CC Pfam: PF00481; PP2C; 1.
 CC SMART: SM00044; CYCC; 1.
 CC SMART: SM00369; LRR_TYP; 1.
 CC SMART: SM00332; PP2CC; 1.

DR SMART; SM00314; RA; 1.
 DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
 DR PROSITE; PS50200; RA; 1.
 KN Lyase, Repeat; Leucine-rich repeat; CAMP biosynthesis; Metal-binding;
 KW Magnesium.
 FT DOMAIN 292 380 RAS-ASSOCIATING.
 FT REPEAT 385 403 LRR 1.
 FT REPEAT 404 427 LRR 2.
 FT REPEAT 429 451 LRR 3.
 FT REPEAT 452 475 LRR 4.
 FT REPEAT 476 498 LRR 5.
 FT REPEAT 501 524 LRR 6.
 FT REPEAT 525 547 LRR 7.
 FT REPEAT 548 570 LRR 8.
 FT REPEAT 572 594 LRR 9.
 FT REPEAT 595 618 LRR 10.
 FT REPEAT 658 681 LRR 11.
 FT REPEAT 682 705 LRR 12.
 FT REPEAT 728 751 LRR 13.
 FT REPEAT 753 776 LRR 14.
 FT REPEAT 781 805 LRR 15.
 FT REPEAT 807 828 LRR 16.
 FT REPEAT 829 852 LRR 17.
 FT REPEAT 854 876 LRR 18.
 FT REPEAT 877 899 LRR 19.
 FT REPEAT 900 922 LRR 20.
 FT REPEAT 928 954 LRR 21.
 FT DOMAIN 1008 1276 PP2C-LIKE.
 FT DOMAIN 1277 1692 CATALYTIC.
 FT METAL 1337 1337 MAGNESIUM (BY SIMILARITY).
 FT METAL 1380 1380 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 1692 AA; 190333 MW; D137CEB8770A8655 CRC64; -

Query Match 66.1%; Score 37; DB 1; Length 1692;

Best Local Similarity 71.4%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NNNWPTT 9

Db 198 NNNWPAS 204

RESULT 7

KV51 MOUSE
 ID KV51 MOUSE STANDARD; PRT; 115 AA.
 AC P01642;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE IG kappa chain V-V region L7 precursor (fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81220975; PubMed=6264318;
 RA Pech M., Hochtl J., Schnell H., Zachau H.G.;
 RT "Differences between germ-line and rearranged immunoglobulin V kappa
 RT coding sequences suggest a localized mutation mechanism.";
 RC Nature 291:668-670(1991).
 CC -!- MISCELLANEOUS: THERE APPEAR TO BE TWO POSSIBLE SPLICE JUNCTIONS AT
 CC THE 3' END OF THE INTRON. THE ALTERNATE WOULD CODE FOR A PROTEIN
 CC LACKING RESIDUES 17-19.
 CC PIR: A01925; KMSL7.
 DR PDB: 1J10; 18-FEB-03.
 DR PDB: 1J1P; 18-FEB-03.
 DR PDB: 1J1X; 18-FEB-03.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.

```

KW Immunoglobulin V region; Signal; 3D-structure.
FT SIGNAL 1 20 IG KAPPA CHAIN V-V REGION L7.
FT CHAIN 21 >115 FRAMEWORK-1.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 >115 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12615 MW; C17BEC758C577E00 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 115;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQNNWP 7
|:|:|
Db 109 QQNSWP 115

RESULT 8
ID_CADN_RAT STANDARD; PRT; 3317 AA.
AC P58365;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cadherin-23 precursor (Otocadherin).
GN CDH23 OR KIAA1774.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=21481445; PubMed=11597768;
RA Nakajima D., Nakayama M., Kikuno R., Hiroseawa M., Nagase T., Ohara O.;
RT "Identification of three novel non-classical cadherin genes through
RT comprehensive analysis of large cDNAs";
RL Brain Res. Mol. Brain Res. 94:85-95(2001).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells. Cadherin 23 may function as hair
CC bundle organizer perhaps by cross-linking the stereocilia (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Contains 27 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB053447; BAB61904.1; -
CC InterPro; IPR002126; Cadherin.
CC Pfam; PF000028; cadherin; 27.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 26.
CC PROSITE; PS00232; CADHERIN 1; 17.
CC PROSITE; PS0268; CADHERIN 2; 27.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 3317 CADHERIN-23.
FT DOMAIN 24 3062 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 3063 3083 POTENTIAL.
FT DOMAIN 3084 3317 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 34 132
FT DOMAIN 133 236
FT DOMAIN 237 348
FT DOMAIN 349 458
FT DOMAIN 459 559
FT DOMAIN 560 669
FT DOMAIN 670 782
FT DOMAIN 777 888
FT DOMAIN 889 993
FT DOMAIN 994 1100
FT DOMAIN 1101 1206
FT DOMAIN 1208 1311
FT DOMAIN 1312 1416
FT DOMAIN 1418 1525
FT DOMAIN 1527 1632
FT DOMAIN 1633 1742
FT DOMAIN 1743 1849
FT DOMAIN 1850 1957
FT DOMAIN 1958 2067
FT DOMAIN 2068 2172
FT DOMAIN 2173 2291
FT DOMAIN 2291 2400
FT DOMAIN 2401 2507
FT DOMAIN 2508 2609
FT DOMAIN 2612 2720
FT DOMAIN 2727 2844
FT DOMAIN 2845 2973
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 349 349 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 600 600 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 825 825 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 999 999 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1169 1169 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1280 1280 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1313 1313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1471 1471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1532 1532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1649 1649 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1665 1665 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1816 1816 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1855 1855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1887 1887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1900 1900 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2012 2012 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2048 2048 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2127 2127 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2166 2166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2193 2193 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2261 2261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2355 2355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2367 2367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2576 2576 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2614 2614 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2747 2747 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2806 2806 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2875 2875 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2894 2894 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2939 2939 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2979 2979 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 3317 AA; 365448 MW; CEP397A066BF0986 CRC64;

```

Query Match 64.3%; Score 36; DB 1; Length 3317;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNNWPT 8
|:||||
Db 2061 NDNWPT 2066

RESULT 9
CADN_MOUSE STANDARD; PRT: 3354 AA.
AC Q99PF4; Q99NH1; Q9D4N9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cadherin-23 precursor (Otocadherin).
GN CDH23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20578758; PubMed=11138008;
RA Di Palma F., Holme R.H., Bryda E.C., Belyantseva I.A., Pellegrino R.,
RA Kachar B., Steel K.P., Noben-Trauth K.;
RT "Mutations in Cdh23, encoding a new type of cadherin, cause
stereocilia disorganization in waltzer, the mouse model for Usher
syndrome type 1D".
RL Nat. Genet. 27:103-107(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=21280917; PubMed=11386759;
RA Wilson S.M., Householder D.B., Coppola V., Tesserollo L., Fritzsche B.,
RA Lee E.-C., Goss D., Carlson G.A., Copeland N.G., Jenkins N.A.;
RT "Mutations in Cdh23 cause nonsyndromic hearing loss in waltzer mice";
RL Genomics 74:228-233(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Fu Y., Wang Q., Roe B.A.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP GENOMIC ORGANIZATION, ALTERNATIVE SPLICING, VARIANT WALTZER
RP 2718-ASU-GLU-PRO-2720 DEL, AND VARIANTS PRO-5; VAL-229; LYS-891;
RP ILE-1137; ARG-1236; VAL-2025; VAL-2026; THR-2217; HIS-2222; ARG-2270
RP AND ALA-2617.
RX MEDLINE=21623040; PubMed=11750125;
RA Di Palma F., Pellegrino R., Noben-Trauth K.;

RT "Genomic structure, alternative splice forms and normal and mutant
RL alleles of cadherin 23 (Cdh23).";
CC Gene 281331-41(2001).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells. Cadherin 23 may function as hair
CC bundle organizer perhaps by cross-linking the stereocilia.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q99PF4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q99PF4-2; Sequence=VSP_000648;
CC -!- TISSUE SPECIFICITY: In adult animals relatively high levels of
CC expression are found in testis, skeletal muscle, heart, eye and
CC thymus, and lower expression in kidney, lung and brain. Found in
CC the sensory hair cells of the inner ear.
CC -!- DISEASE: Defects in CDH23 are the cause of waltzer (v) phenotype.
CC Waltzer mice are characterized by deafness and vestibular
CC dysfunction due to degeneration of the neuroepithelium within the
CC inner ear.
CC -!- SIMILARITY: Contains 27 cadherin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF308939; AAG52817.1; -;
CC EMBL; AY026062; AAK07670.1; -;
CC EMBL; AK016365; -; NOT_ANNOTATED_CDS.
CC EMBL; AC079818; -; NOT_ANNOTATED_CDS.
CC EMBL; AC079819; -; NOT_ANNOTATED_CDS.
CC EMBL; AC079082; -; NOT_ANNOTATED_CDS.
CC HSP; F15116; INCU.
CC DR MGD; MGI:1890219; Cdh23.
CC GO; GO:0005929; C:Cellium; IDA.
CC GO; GO:0005515; F:Protein binding; IPI.
CC InterPro; IPR002126; Cadherin.
CC Pfam; PF00028; cadherin; 27.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 26.
CC DR PROSITE; PS00232; CADHERIN_1; 17.
CC DR PROSITE; PS00268; CADHERIN_2; 27.
CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
CC Signal; Multigene family; Alternative splicing; Deafness;
CC Polymorphism; Disease mutation
CC SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 3354 CADHERIN-23.
FT DOMAIN 24 3064 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 3065 3085 POTENTIAL.
FT DOMAIN 3086 3354 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 132 CADHERIN_1.
FT DOMAIN 133 236 CADHERIN_2.
FT DOMAIN 237 348 CADHERIN_3.
FT DOMAIN 349 460 CADHERIN_4.
FT DOMAIN 461 561 CADHERIN_5.
FT DOMAIN 562 671 CADHERIN_6.
FT DOMAIN 672 784 CADHERIN_7.
FT DOMAIN 779 890 CADHERIN_8.
FT DOMAIN 891 995 CADHERIN_9.
FT DOMAIN 996 1102 CADHERIN_10.
FT DOMAIN 1103 1208 CADHERIN_11.
FT DOMAIN 1210 1313 CADHERIN_12.
FT DOMAIN 1314 1418 CADHERIN_13.
FT DOMAIN 1420 1527 CADHERIN_14.
FT DOMAIN 1529 1634 CADHERIN_15.
FT DOMAIN 1635 1744 CADHERIN_16.
FT DOMAIN 1745 1851 CADHERIN_17.

```
FT DOMAIN 1852 1959 CADHERIN 18.
FT DOMAIN 1960 2069 CADHERIN 19.
FT DOMAIN 2070 2174 CADHERIN 20.
FT DOMAIN 2175 2293 CADHERIN 21.
FT DOMAIN 2297 2402 CADHERIN 22.
FT DOMAIN 2403 2509 CADHERIN 23.
FT DOMAIN 2510 2611 CADHERIN 24.
FT DOMAIN 2614 2722 CADHERIN 25.
FT DOMAIN 2729 2846 CADHERIN 26.
FT DOMAIN 2847 2975 CADHERIN 27.
FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 349 349 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 434 434 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 472 472 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 694 694 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 765 765 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 810 810 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 941 941 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1001 1001 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1171 1171 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1282 1282 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1315 1315 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1473 1473 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1534 1534 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1651 1651 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1667 1667 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1818 1818 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1857 1857 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1889 1889 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 1902 1902 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2014 2014 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2050 2050 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2129 2129 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2168 2168 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2195 2195 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2263 2263 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2357 2357 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2369 2369 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2578 2578 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2616 2616 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2749 2749 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2808 2808 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2877 2877 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2896 2896 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2941 2941 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT CARBOHYD 2981 2981 N-LINKED (GLCNAC. .) (POTENTIAL. .)
FT VARSPLIC 3212 3246 Missing (in isoform 2).
FT VARIANT 5 5 /FTidvSP 000648.
FT VARIANT 229 229 L -> P (IN STRAIN CAST/EI).
FT VARIANT 891 891 M -> V (IN STRAIN CAST/EI).
FT VARIANT 1137 1137 R -> K (IN STRAIN CAST/EI).
FT VARIANT 1236 1236 V -> I (IN STRAIN CAST/EI).
FT VARIANT 2025 2025 K -> R (IN STRAIN CAST/EI).
FT VARIANT 2026 2026 I -> V (IN STRAIN CAST/EI).
FT VARIANT 2217 2217 I -> T (IN STRAIN CAST/EI).
FT VARIANT 2222 2222 R -> H (IN STRAIN CAST/EI).

Query Match 64.3%; Score 36; DB 1; Length 3354;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNNWPT 8
Db 2063 NDNWPT 2068
|:|:|:|:|
|:|:|:|:|
```

FT NP_BIND 315 322 ATP (POTENTIAL).
 FT NP_BIND 653 660 ATP (POTENTIAL).
 FT NP_BIND 1083 1090 ATP (POTENTIAL).
 FT NP_BIND 1368 1375 ATP (POTENTIAL).
 FT NP_BIND 1747 1754 ATP (POTENTIAL).
 FT NP_BIND 2054 2061 ATP (POTENTIAL).
 FT NP_BIND 2681 2688 POLY-LYS.
 FT NP_BIND 2768 2771 POLY-LYS.
 FT DOMAIN 2904 2907 POLY-LEU.
 FT DOMAIN 4136 4139 POLY-LEU.
 FT DOMAIN 4704 4899 VMFA.
 SQ SEQUENCE 4910 AA; 559302 MW; E4E873BEDF6E1E5B CRC64;

Query Match 64.3%; Score 36; DB 1; Length 4910;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NNNWP 7
 Db 2045 NNNWP 2049

RESULT 11
 ID KV31 HUMAN STANDARD; PRT; 115 AA.
 AC P04433;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG kappa chain V-III region VG precursor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85087932; PubMed=6440122;
 RA Pech M., Zachau H.G.;
 RT "Immunoglobulin genes of different subgroups are interdigitated
 within the VK locus.";
 RL Nucleic Acids Res. 12:9229-9236 (1984).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; A01668; ; NOT_ANNOTATED_CDS.
 DR PIR; A01900; K3HUVG.
 DR HSSP; P80362; 1WTL.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0003823; F:antigen binding; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; Ig; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Immunoglobulin V region; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 >115 IG KAPPA CHAIN V-III REGION VG.
 FT DOMAIN 21 43 FRAMEWORK-1.
 FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 55 59 FRAMEWORK-2.
 FT DOMAIN 60 76 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 77 108 FRAMEWORK-3.
 FT DOMAIN 109 115 COMPLEMENTARITY-DETERMINING-3.
 FT DISULFID 43 108 BY SIMILARITY.
 FT NON_TER 115 115
 SQ SEQUENCE 115 AA; 12575 MW; 2DE47CDA3A17D555 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 115;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNWP 7
 Db 109 QQSNWP 115

RESULT 12
 ID RADC_XYLFA STANDARD; PRT; 224 AA.
 AC Q9PGZ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA repair protein radC homolog.
 GN RADC OR XF0148.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinsch F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neco C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Dosena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Paixoto B.B., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159 (2000).
 CC -!- FUNCTION: Involved in DNA repair (by similarity).
 CC -!- SIMILARITY: Belongs to the radC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; A5003868; AAF82961.1; ALT_INIT.
 DR HAWAF; NF_00018; ; 1.
 DR InterPro; IPR001405; RadC.
 DR Pfam; PF04002; RadC; 1.
 DR ProDom; PD007415; RadC; 1.
 DR TIGRFAMs; TIGR00608; radC; 1.
 DR PROSITE; PS01302; RADC; 1.
 DR DNA repair; Complete proteome.

SQ SEQUENCE 224 AA; 24618 MW; 474351E63C890A5A CRC64;
Query Match 62.5%; Score 35; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 NNWPT 8
| | | | |
Db 4 NNWPT 8
RESULT 13
RADCL XYLFT STANDARD; PRT; 224 AA.
AC Q87F21;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA repair protein radC homolog.
GN RADCL OR PD0117.
OS Xylella fastidiosa (strain Temecul / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
ON NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Issai S.M.,
RA Carlier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.L., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa";
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- FUNCTION: Involved in DNA repair (By similarity).
CC -!- SIMILARITY: Belongs to the radC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012553; AAC28016.1; -
DR HAMAP; MF_00018; -; 1.
DR InterPro; IPR001405; RadC.
DR Pfam; PF04002; RadC; 1.
DR PROSITE; PS01302; RADCL; 1.
KW DNA repair; Complete proteome.
SQ SEQUENCE 224 AA; 24590 MW; B84351B63C8AF775 CRC64;
Query Match 62.5%; Score 35; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 NNWPT 8
| | | | |
Db 4 NNWPT 8
RESULT 14

PROM ECOLI
ID PROM ECOLI STANDARD; PRT; 354 AA.
AC P14176;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycine betaine/L-proline transport system permease protein prom.
GN PROM OR B2678.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89197759; PubMed=2649479;
RA Gowrishankar J.;
RT "Nucleotide sequence of the osmoregulatory proU operon of Escherichia
RT coli";
RL J. Bacteriol. 171:1923-1931(1989).
RN [2]
RP ERRATUM.
RA Gowrishankar J.;
RL J. Bacteriol. 172:1165-1165(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=X12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=X12;
RC MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshibuchi T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113(1997).
CC -!- FUNCTION: INVOLVED IN A MULTICOMPONENT BINDING-PROTEIN-DEPENDENT
CC TRANSPORT SYSTEM FOR GLYCINE BETAINE/L-PROLINE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family. CysTW subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M24856; AAA24428.1; -
DR EMBL; AE000352; AAC75725.1; -
DR EMBL; D90891; BAA16543.1; -
DR FIC; J00129; MMECFW.
DR EcoGene; EG10772; prom.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
DR PROSITE; PS05028; ABC_TM1; 1.
KW Transport; Amino-acid-transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT DOMAIN 1 99 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 100 120 POTENTIAL.

FT DOMAIN 121 121 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 122 142 POTENTIAL.
 FT DOMAIN 143 150 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 151 171 POTENTIAL.
 FT DOMAIN 172 198 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 199 219 POTENTIAL.
 FT DOMAIN 220 270 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 271 291 POTENTIAL.
 FT DOMAIN 292 300 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 301 321 POTENTIAL.
 FT DOMAIN 322 354 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 354 AA; 37619 MW; D35F9A74E2779D1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 354;
 Best Local Similarity 75.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 QNNWPPT 9
 DB 4 QNNPDDT 11

RESULT 15
 MALQ_CHLTR STANDARD; PRT; 527 AA.
 AC O84059;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
 DE (Disproportionating enzyme) (D-enzyme).
 GN MALQ OR CT087.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/UW-3/CX;
 RX MEDLINE=99000609; PubMed=9784136;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis";
 RL Science 282:754-759(1998).
 CC -!- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
 CC to a new 4-position in an acceptor, which may be glucose or (1,4)-
 CC alpha-D-glucan.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the disproportionating enzyme family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF001283; AAC67678.1; -.
 CC FIC; G71557; G71557.
 CC InterPro: IPR003385; Glycohydro.77.
 CC Pfam: PF02446; 4A_Glucanotrans; 1.
 CC TIGRFAMs: TIGR00217; malQ; 1.
 KW Transferase; Glycosyltransferase; Carbohydrate metabolism;
 KW Complete proteome.
 SQ SEQUENCE 527 AA; 61453 MW; 52455D5E9ED2D46D CRC64;

Query Match 62.5%; Score 35; DB 1; Length 527;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 NNNPPT 9

Db 186 NNNPPT 191
 Search completed: October 6, 2004, 16:30:32
 Job time : 7.52632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:34:15 ; Search time 56.0526 Seconds
(without alignments)
51.669 Million cell updates/sec

Title: US-09-635-974A-12

Perfect score: 56

Sequence: 1 QQNNWPTT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 56 | 100.0 | 9 | 10 | US-09-798-689-36 |
| 2 | 56 | 100.0 | 9 | 10 | US-09-996-954B-12 |
| 3 | 56 | 100.0 | 9 | 12 | US-10-374-600-112 |
| 4 | 56 | 100.0 | 9 | 15 | US-10-374-531-112 |
| 5 | 56 | 100.0 | 107 | 12 | US-10-374-600-114 |
| 6 | 56 | 100.0 | 107 | 15 | US-10-374-531-114 |
| 7 | 56 | 100.0 | 113 | 12 | US-10-374-600-20 |
| 8 | 56 | 100.0 | 113 | 12 | US-10-374-600-22 |
| 9 | 56 | 100.0 | 113 | 12 | US-10-374-600-23 |
| 10 | 56 | 100.0 | 113 | 15 | US-10-374-531-20 |
| 11 | 56 | 100.0 | 113 | 15 | US-10-374-531-22 |
| 12 | 56 | 100.0 | 113 | 15 | US-10-374-531-23 |
| 13 | 56 | 100.0 | 127 | 12 | US-10-374-600-5 |
| 14 | 56 | 100.0 | 127 | 12 | US-10-374-600-11 |
| 15 | 56 | 100.0 | 127 | 12 | US-10-374-600-15 |

| | | | | | |
|----|----|-------|------|----|--------------------|
| 16 | 56 | 100.0 | 127 | 12 | US-10-374-600-17 |
| 17 | 56 | 100.0 | 127 | 15 | US-10-374-531-5 |
| 18 | 56 | 100.0 | 127 | 15 | US-10-374-531-11 |
| 19 | 56 | 100.0 | 127 | 15 | US-10-374-531-15 |
| 20 | 56 | 100.0 | 127 | 15 | US-10-374-531-17 |
| 21 | 50 | 89.3 | 107 | 12 | US-10-374-600-113 |
| 22 | 50 | 89.3 | 107 | 15 | US-10-374-531-113 |
| 23 | 45 | 80.4 | 106 | 12 | US-10-453-698-27 |
| 24 | 45 | 80.4 | 106 | 15 | US-10-308-817-27 |
| 25 | 45 | 80.4 | 107 | 12 | US-10-453-698-129 |
| 26 | 45 | 80.4 | 107 | 12 | US-10-453-698-130 |
| 27 | 45 | 80.4 | 107 | 15 | US-10-308-817-129 |
| 28 | 45 | 80.4 | 107 | 15 | US-10-308-817-130 |
| 29 | 45 | 80.4 | 108 | 15 | US-10-412-703A-129 |
| 30 | 44 | 78.6 | 224 | 10 | US-09-453-234-64 |
| 31 | 43 | 76.8 | 107 | 12 | US-10-251-085B-123 |
| 32 | 43 | 76.8 | 107 | 16 | US-10-737-252-123 |
| 33 | 43 | 76.8 | 108 | 12 | US-10-251-085B-132 |
| 34 | 43 | 76.8 | 108 | 16 | US-10-737-252-132 |
| 35 | 43 | 76.8 | 1558 | 15 | US-10-080-334-230 |
| 36 | 42 | 75.0 | 107 | 12 | US-10-389-417-47 |
| 37 | 42 | 75.0 | 107 | 12 | US-10-452-357-34 |
| 38 | 42 | 75.0 | 107 | 12 | US-10-251-085B-108 |
| 39 | 42 | 75.0 | 107 | 12 | US-10-251-085B-109 |
| 40 | 42 | 75.0 | 107 | 12 | US-10-251-085B-117 |
| 41 | 42 | 75.0 | 107 | 12 | US-10-251-085B-119 |
| 42 | 42 | 75.0 | 107 | 12 | US-10-251-085B-121 |
| 43 | 42 | 75.0 | 107 | 12 | US-10-251-085B-122 |
| 44 | 42 | 75.0 | 107 | 12 | US-10-251-085B-125 |
| 45 | 42 | 75.0 | 107 | 14 | US-10-330-613-6 |

ALIGNMENTS

RESULT 1

US-09-798-689-36
; Sequence 36, Application US/09798689
; Publication No. US20030103973A1
; GENERAL INFORMATION:
; APPLICANT: Rockwell, Patricia
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: Method for Reducing Tumor Growth with VEGF Antagonists
; TITLE OF INVENTION: Combined With Radiation and Chemotherapy
; FILE REFERENCE: Sequence Listings 1-41 for 381-25 CIP
; CURRENT APPLICATION NUMBER: US/09/798,689
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/401,163
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: 08/967,113
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 08/706,804
; PRIOR FILING DATE: 1996-09-03
; PRIOR APPLICATION NUMBER: 08/476,533
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/326,552
; PRIOR FILING DATE: 1994-10-20
; PRIOR APPLICATION NUMBER: 08/196,041
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Mouse
US-09-798-689-36

Query Match 100.0%; Score 56; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQNNWPTT 9
|||||||

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-10-374-531-112

Query Match 100.0%; Score 56; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNNWPTT 9
Db 1 QQNNNWPTT 9

RESULT 5
US-10-374-600-114
; Sequence 114, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage

COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-7200
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 114:

US-10-374-600-114
Query Match 100.0%; Score 56; DB 12; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNNWPTT 9

Db 89 QQNNNWPTT 97

RESULT 6
US-10-374-531-114
; Sequence 114, Application US/10374531
; Publication No. US20040006212A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELEPHONE: (212) 425-7200
TELEFAX: (212) 425-5288
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 114:

US-10-374-531-114
Query Match 100.0%; Score 56; DB 15; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNNWPTT 9
Db 89 QQNNNWPTT 97

RESULT 7
US-10-374-600-20
; Sequence 20, Application US/10374600
; Publication No. US20030224001A1
; GENERAL INFORMATION:
; APPLICANT: ImClone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR

INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-10-374-600-20

Query Match 100.0%; Score 56; DB 12; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 QQNNNWPTT 9
Db 95 QQNNNWPTT 103

RESULT 8

US-10-374-600-22
Sequence 22, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847

COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 22:

US-10-374-600-22

Query Match 100.0%; Score 56; DB 12; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 QQNNNWPTT 9
Db 95 QQNNNWPTT 103

RESULT 9

US-10-374-600-23
Sequence 23, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:
APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847

```
/
/ FILING DATE: 07-JUN-1996
/ APPLICATION NUMBER: US 08/482,982
/ FILING DATE: 07-JUN-1995
/ APPLICATION NUMBER: US 08/573,289
/ FILING DATE: 15-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Deborah A. Somerville
/ REGISTRATION NUMBER: 31,995
/ REFERENCE/DOCKET NUMBER: 11245/46003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 425-5288
/ TELEFAX: (212) 425-5288
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 113 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-374-600-23

Query Match 100.0%; Score 56; DB 12; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 QQNNNWPTT 9
Db 95 QQNNNWPTT 103

RESULT 10
US-10-374-531-20
/ Sequence 20, Application US/10374531
/ Publication No. US20040006212A1
/ GENERAL INFORMATION:
/ APPLICANT: ImClone Systems Incorporated, et al.
/ TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
/ INHIBITING THE GROWTH OF TUMORS
/ NUMBER OF SEQUENCES: 120
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Kenyon & Kenyon
/ STREET: One Broadway
/ CITY: New York
/ STATE: New York
/ COUNTRY: US
/ ZIP: 10004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/374,531
/ FILING DATE: 25-Feb-2003
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/973,065C
/ FILING DATE: 19-Mar-1998
/ APPLICATION NUMBER: PCT/US96/09847
/ FILING DATE: 07-JUN-1996
/ APPLICATION NUMBER: US 08/482,982
/ FILING DATE: 07-JUN-1995
/ APPLICATION NUMBER: US 08/573,289
/ FILING DATE: 15-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Deborah A. Somerville
/ REGISTRATION NUMBER: 31,995
/ REFERENCE/DOCKET NUMBER: 11245/46003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 425-7200
```

```
/
/ TELEFAX: (212) 425-5288
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 113 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-374-531-20

Query Match 100.0%; Score 56; DB 15; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNNWPTT 9
Db 95 QQNNNWPTT 103

RESULT 11
US-10-374-531-22
/ Sequence 22, Application US/10374531
/ Publication No. US20040006212A1
/ GENERAL INFORMATION:
/ APPLICANT: ImClone Systems Incorporated, et al.
/ TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
/ INHIBITING THE GROWTH OF TUMORS
/ NUMBER OF SEQUENCES: 120
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Kenyon & Kenyon
/ STREET: One Broadway
/ CITY: New York
/ STATE: New York
/ COUNTRY: US
/ ZIP: 10004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/374,531
/ FILING DATE: 25-Feb-2003
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/973,065C
/ FILING DATE: 19-Mar-1998
/ APPLICATION NUMBER: PCT/US96/09847
/ FILING DATE: 07-JUN-1996
/ APPLICATION NUMBER: US 08/482,982
/ FILING DATE: 07-JUN-1995
/ APPLICATION NUMBER: US 08/573,289
/ FILING DATE: 15-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Deborah A. Somerville
/ REGISTRATION NUMBER: 31,995
/ REFERENCE/DOCKET NUMBER: 11245/46003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 425-7200
/ TELEFAX: (212) 425-5288
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 113 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FRAGMENT TYPE: internal
/ SEQUENCE DESCRIPTION: SEQ ID NO: 22:
```

US-10-374-531-22

Query Match 100.0%; Score 56; DB 15; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9
|||
Db 95 QONNNWPTT 103

RESULT 12

US-10-374-531-23
Sequence 23, Application US/10374531
Publication No. US2004006212A1
GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,531
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 23:

US-10-374-531-23
Query Match 100.0%; Score 56; DB 15; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9
|||
Db 95 QONNNWPTT 103

RESULT 13

US-10-374-600-5
Sequence 5, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: One Broadway
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/374,600
FILING DATE: 25-Feb-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/973,065C
FILING DATE: 19-Mar-1998
APPLICATION NUMBER: PCT/US96/09847
FILING DATE: 07-JUN-1996
APPLICATION NUMBER: US 08/482,982
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/573,289
FILING DATE: 15-DEC-1995

ATTORNEY/AGENT INFORMATION:
NAME: Deborah A. Somerville
REGISTRATION NUMBER: 31,995
REFERENCE/DOCKET NUMBER: 11245/46003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 425-5288
TELEFAX: (212) 425-5288

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-10-374-600-5
Query Match 100.0%; Score 56; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9
|||
Db 109 QONNNWPTT 117

RESULT 14

US-10-374-600-11
Sequence 11, Application US/10374600
Publication No. US20030224001A1
GENERAL INFORMATION:

APPLICANT: Imclone Systems Incorporated, et al.
TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
INHIBITING THE GROWTH OF TUMORS

NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon

```

; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-374-600-11
;
; Query Match 100.0%; Score 56; DB 12; Length 127;
; Best Local Similarity 100.0%; Pred. No. 0.21;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 QQNNNWPTT 9
; DB 109 QQNNNWPTT 117
;
; RESULT 15
; US-10-374-600-15
; Sequence 15, Application US/10374600
; Publication No. US20030224001A1
;
; GENERAL INFORMATION:
; APPLICANT: Imclone Systems Incorporated, et al.
; TITLE OF INVENTION: ANTIBODY AND ANTIBODY FRAGMENTS FOR
; INHIBITING THE GROWTH OF TUMORS
;
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: One Broadway
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
;
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/10/374,600
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,065C
; FILING DATE: 19-Mar-1998
; APPLICATION NUMBER: PCT/US96/09847
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 08/482,982
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/573,289
; FILING DATE: 15-DEC-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Deborah A. Somerville
; REGISTRATION NUMBER: 31,995
; REFERENCE/DOCKET NUMBER: 11245/46003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 425-7200
; TELEFAX: (212) 425-5288
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-374-600-15
;
; Query Match 100.0%; Score 56; DB 12; Length 127;
; Best Local Similarity 100.0%; Pred. No. 0.21;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 QQNNNWPTT 9
; DB 109 QQNNNWPTT 117
;
; Search completed: October 6, 2004, 17:09:06
; Job time : 57.0526 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:24:54 ; Search time 15.9474 Seconds
(without alignments)
29.135 Million cell updates/sec

Title: US-09-635-974A-12
Perfect score: 56
Sequence: 1 QONNNWPTT 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCUTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 45 | 80.4 | 240 | 2 | US-07-956-399-2 |
| 2 | 44 | 78.6 | 224 | 4 | US-09-456-090A-84 |
| 3 | 44 | 78.6 | 239 | 2 | US-07-956-399-4 |
| 4 | 42 | 75.0 | 107 | 1 | US-07-634-278-34 |
| 5 | 42 | 75.0 | 107 | 1 | US-08-477-728-34 |
| 6 | 42 | 75.0 | 107 | 1 | US-08-474-040-34 |
| 7 | 42 | 75.0 | 107 | 1 | US-08-487-200-34 |
| 8 | 42 | 75.0 | 107 | 3 | US-08-484-537-34 |
| 9 | 41 | 73.2 | 9 | 2 | US-08-476-176B-55 |
| 10 | 41 | 73.2 | 9 | 3 | US-08-127-721A-55 |
| 11 | 41 | 73.2 | 9 | 3 | US-08-485-246A-55 |
| 12 | 41 | 73.2 | 106 | 2 | US-08-800-198-4 |
| 13 | 41 | 73.2 | 106 | 3 | US-09-296-595-4 |
| 14 | 41 | 73.2 | 107 | 2 | US-08-476-176B-4 |
| 15 | 41 | 73.2 | 107 | 3 | US-08-127-721A-4 |
| 16 | 41 | 73.2 | 107 | 3 | US-08-485-246A-4 |
| 17 | 41 | 73.2 | 127 | 2 | US-08-476-176B-6 |
| 18 | 41 | 73.2 | 127 | 2 | US-08-476-176B-8 |
| 19 | 41 | 73.2 | 127 | 2 | US-08-476-176B-10 |
| 20 | 41 | 73.2 | 127 | 3 | US-08-127-721A-6 |
| 21 | 41 | 73.2 | 127 | 3 | US-08-127-721A-8 |
| 22 | 41 | 73.2 | 127 | 3 | US-08-127-721A-10 |
| 23 | 41 | 73.2 | 127 | 3 | US-08-485-246A-6 |
| 24 | 41 | 73.2 | 127 | 3 | US-08-485-246A-8 |
| 25 | 41 | 73.2 | 127 | 3 | US-08-485-246A-10 |
| 26 | 41 | 73.2 | 240 | 2 | US-08-800-198-8 |
| 27 | 41 | 73.2 | 240 | 3 | US-09-296-595-8 |

28 40 71.4 103 1 US-08-436-463-21 Sequence 21, Appl
29 40 71.4 127 1 US-08-436-463-18 Sequence 18, Appl
30 40 71.4 244 4 US-09-244-369B-1 Sequence 1, Appl
31 39 69.6 9 4 US-09-339-922A-86 Sequence 86, Appl
32 39 69.6 9 4 US-09-016-061-86 Sequence 86, Appl
33 39 69.6 33 4 US-08-525-539A-12 Sequence 12, Appl
34 39 69.6 100 1 US-08-436-463-19 Sequence 19, Appl
35 39 69.6 107 1 US-07-634-278-62 Sequence 62, Appl
36 39 69.6 107 1 US-07-634-278-63 Sequence 63, Appl
37 39 69.6 107 1 US-07-634-278-87 Sequence 87, Appl
38 39 69.6 107 1 US-08-477-728-62 Sequence 62, Appl
39 39 69.6 107 1 US-08-477-728-63 Sequence 63, Appl
40 39 69.6 107 1 US-08-477-728-87 Sequence 87, Appl
41 39 69.6 107 1 US-08-474-040-62 Sequence 62, Appl
42 39 69.6 107 1 US-08-474-040-63 Sequence 63, Appl
43 39 69.6 107 1 US-08-474-040-87 Sequence 87, Appl
44 39 69.6 107 1 US-08-487-200-62 Sequence 62, Appl
45 39 69.6 107 1 US-08-487-200-63 Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-07-956-399-2
; Sequence 2, Application US/07956399
; Patent No. 5876717
; GENERAL INFORMATION:
; APPLICANT: SHIMAMURA, TOSHIRO
; APPLICANT: TAKI, SHINSUKE
; APPLICANT: HAMURO, JUNJI
; TITLE OF INVENTION: POLYPEPTIDES CAPABLE OF BINDING TO HEAVY
; TITLE OF INVENTION: CHAINS OF IL-2 RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,399
; FILING DATE: 19921005
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5876717man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-586-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-956-399-2

Query Match 80.4%; Score 45; DB 2; Length 240;
Best Local Similarity 77.8%; Pred. No. 6.6;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 QONNNWPTT 9
|||:||||

```
Db          90 QQTNSWPTT 98
RESULT 2
US-09-456-090A-84
; Sequence 84, Application US/09456090A
; Patent No. 6680209
; GENERAL INFORMATION:
; APPLICANT: Buechler, Joe
; APPLICANT: Walkiers, Gunars
; APPLICANT: Gray, Jeff
; APPLICANT: Lomberg, Nils
; FILE OF INVENTION: HUMAN ANTIBODIES AS DIAGNOSTIC REAGENTS
; FILE REFERENCE: 020015-000200US
; CURRENT APPLICATION NUMBER: US/09/456,090A
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 84
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: M2-32L
US-09-456-090A-84

Query Match      78.6%; Score 44; DB 4; Length 224;
Best Local Similarity 77.8%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY          1 QQNNWPTT 9
          |||||
          89 QQNNWPLT 97

Db          90 QQTNSWPTT 98
RESULT 3
US-09-456-399-4
; Sequence 4, Application US/07956399
; Patent No. 5876717
; GENERAL INFORMATION:
; APPLICANT: SHIMAMURA, TOSHIRO
; APPLICANT: TAKI, SHINSUKE
; APPLICANT: HAMURO, JUNJI
; TITLE OF INVENTION: POLYPEPTIDES CAPABLE OF BINDING TO HEAVY
; TITLE OF INVENTION: CHAINS OF IL-2 RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1765 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,399
; FILING DATE: 19921005
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5876717/man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-586-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids

Db          90 QQTNSWPTT 98
RESULT 4
US-07-634-278-34
; Sequence 34, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COBLEIGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION/DOCKET NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..107
; OTHER INFORMATION: /note= "amino acid sequence of the
; OTHER INFORMATION: light chain for humane Lay antibody."
US-07-634-278-34
```

Query Match 75.0%; Score 42; DB 1; Length 107;
Best Local Similarity 77.8%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQNNWPPT 9
DB 89 QQYNNWPPT 97

RESULT 5
US-08-477-728-34
; Sequence 34, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,728
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..107
; OTHER INFORMATION: /note= "Amino acid sequence of the
; OTHER INFORMATION: light chain for humane Lay antibody."
US-08-477-728-34

Query Match 75.0%; Score 42; DB 1; Length 107;
Best Local Similarity 77.8%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQNNWPPT 9
DB 89 QQYNNWPPT 97

RESULT 6
US-08-474-040-34
; Sequence 34, Application US/08474040
; Patent No. 5693761
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,040
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/634,278
; FILING DATE: 19-DEC-1990
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..107
; OTHER INFORMATION: /note= "Amino acid sequence of the
; OTHER INFORMATION: light chain for humane Lay antibody."
US-08-474-040-34

Query Match 75.0%; Score 42; DB 1; Length 107;
Best Local Similarity 77.8%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQNNWPPT 9

| | | | |
|--|-----------------|--|-----------------|
| Db | 89 QQYNNWPPT 97 | Db | 89 QQYNNWPPT 97 |
| RESULT 7 | | RESULT 8 | |
| US-08-487-200-34 | | US-08-484-537-34 | |
| ; Sequence 34, Application US/08487200 | | ; Sequence 34, Application US/08484537 | |
| ; Patent No. 5693762 | | ; Patent No. 6180370 | |
| ; GENERAL INFORMATION: | | ; GENERAL INFORMATION: | |
| ; APPLICANT: QUEEN, Cary L. | | ; APPLICANT: QUEEN, Cary L. | |
| ; APPLICANT: CO, Man Sung | | ; APPLICANT: CO, Man Sung | |
| ; APPLICANT: SCHNEIDER, William P. | | ; APPLICANT: SCHNEIDER, William P. | |
| ; APPLICANT: LANDOLFI, Nicholas P. | | ; APPLICANT: LANDOLFI, Nicholas P. | |
| ; APPLICANT: COELINGH, Kathleen L. | | ; APPLICANT: COELINGH, Kathleen L. | |
| ; APPLICANT: SELICK, Harold E. | | ; APPLICANT: SELICK, Harold E. | |
| ; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS | | ; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS | |
| ; NUMBER OF SEQUENCES: 113 | | ; NUMBER OF SEQUENCES: 113 | |
| ; CORRESPONDENCE ADDRESS: | | ; CORRESPONDENCE ADDRESS: | |
| ; ADDRESSEE: Townsend and Townsend and Crew | | ; ADDRESSEE: Townsend and Townsend Khourie and Crew | |
| ; STREET: 379 Lytton Avenue | | ; STREET: 379 Lytton Avenue | |
| ; CITY: Palo Alto | | ; CITY: Palo Alto | |
| ; STATE: California | | ; STATE: California | |
| ; COUNTRY: US | | ; COUNTRY: US | |
| ; ZIP: 94301 | | ; ZIP: 94301 | |
| ; COMPUTER READABLE FORM: | | ; COMPUTER READABLE FORM: | |
| ; MEDIUM TYPE: Floppy disk | | ; MEDIUM TYPE: Floppy disk | |
| ; COMPUTER: IBM PC compatible | | ; COMPUTER: IBM PC compatible | |
| ; OPERATING SYSTEM: PC-DOS/MS-DOS | | ; OPERATING SYSTEM: PC-DOS/MS-DOS | |
| ; SOFTWARE: Patent In Release #1.0, Version #1.25 | | ; SOFTWARE: Patent In Release #1.0, Version #1.25 | |
| ; CURRENT APPLICATION DATA: | | ; CURRENT APPLICATION DATA: | |
| ; APPLICATION NUMBER: US/08/487,200 | | ; APPLICATION NUMBER: US/08/484,537 | |
| ; FILING DATE: 7-JUN-1995 | | ; FILING DATE: 28-SEP-1990 | |
| ; CLASSIFICATION: 424 | | ; CLASSIFICATION: | |
| ; PRIOR APPLICATION DATA: | | ; PRIOR APPLICATION DATA: | |
| ; APPLICATION NUMBER: US 07/634,278 | | ; APPLICATION NUMBER: US 07/634,278 | |
| ; FILING DATE: 19-DEC-1990 | | ; FILING DATE: 19-DEC-1990 | |
| ; PRIOR APPLICATION DATA: | | ; PRIOR APPLICATION DATA: | |
| ; APPLICATION NUMBER: US 07/590,274 | | ; APPLICATION NUMBER: US 07/590,274 | |
| ; FILING DATE: 28-SEP-1990 | | ; FILING DATE: 28-SEP-1990 | |
| ; PRIOR APPLICATION DATA: | | ; PRIOR APPLICATION DATA: | |
| ; APPLICATION NUMBER: US 07/310,252 | | ; APPLICATION NUMBER: US 07/310,252 | |
| ; FILING DATE: 13-FEB-1989 | | ; FILING DATE: 13-FEB-1989 | |
| ; PRIOR APPLICATION DATA: | | ; PRIOR APPLICATION DATA: | |
| ; APPLICATION NUMBER: US 07/290,975 | | ; APPLICATION NUMBER: US 07/290,975 | |
| ; FILING DATE: 28-DEC-1988 | | ; FILING DATE: 28-DEC-1988 | |
| ; ATTORNEY/AGENT INFORMATION: | | ; ATTORNEY/AGENT INFORMATION: | |
| ; NAME: Smith, William M | | ; NAME: Smith, William M | |
| ; REGISTRATION NUMBER: 30,223 | | ; REGISTRATION NUMBER: 30,223 | |
| ; REFERENCE/DOCKET NUMBER: 11823-002610 | | ; REFERENCE/DOCKET NUMBER: 11823-002600 | |
| ; TELEPHONE: (415) 326-2400 | | ; TELEPHONE: (415) 326-2400 | |
| ; TELEFAX: (415) 326-2422 | | ; TELEFAX: (415) 326-2422 | |
| ; INFORMATION FOR SEQ ID NO: 34: | | ; INFORMATION FOR SEQ ID NO: 34: | |
| ; SEQUENCE CHARACTERISTICS: | | ; SEQUENCE CHARACTERISTICS: | |
| ; LENGTH: 107 amino acids | | ; LENGTH: 107 amino acids | |
| ; TYPE: amino acid | | ; TYPE: amino acid | |
| ; STRANDEDNESS: single | | ; STRANDEDNESS: single | |
| ; TOPOLOGY: unknown | | ; TOPOLOGY: unknown | |
| ; MOLECULE TYPE: protein | | ; MOLECULE TYPE: protein | |
| ; HYPOTHETICAL: NO | | ; HYPOTHETICAL: NO | |
| ; FEATURE: | | ; FEATURE: | |
| ; NAME/KEY: Protein | | ; NAME/KEY: Protein | |
| ; LOCATION: 1..107 | | ; LOCATION: 1..107 | |
| ; OTHER INFORMATION: /note= "Amino acid sequence of the | | ; OTHER INFORMATION: /note= "Amino acid sequence of the | |
| ; OTHER INFORMATION: light chain for humane Lay antibody." | | ; OTHER INFORMATION: light chain for humane Lay antibody." | |
| US-08-487-200-34 | | US-08-484-537-34 | |
| Query Match | | Query Match | |
| Best Local Similarity | | Best Local Similarity | |
| Matches | | Matches | |
| 7; Conservative | 0; Mismatches | 7; Conservative | 0; Mismatches |
| 2; Indels | 0; Gaps | 2; Indels | 0; Gaps |
| 0; | | 0; | |
| 1 QQYNNWPPT 9 | | 1 QQYNNWPPT 9 | |
| Db | 89 QQYNNWPPT 97 | Db | 89 QQYNNWPPT 97 |

```

RESULT 9
US-08-476-176B-55
; Sequence 55, Application US/08476176B
; Patent No. 5958708
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 5958708man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066718artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,176B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1993
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-176B-55

Query Match 73.2%; Score 41; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNWPPT 9
Db 1 QQSDSWPTT 9

RESULT 10
US-08-127-721A-55
; Sequence 55, Application US/08127721A
; Patent No. 6066718
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6066718man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066718artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,246A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; NAME: No. 6072035ak, Henry P.
; REGISTRATION NUMBER: 33,200
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-127-721A-55

Query Match 73.2%; Score 41; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,721A
; FILING DATE: 27-SEPTEMBER-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6066718ak, Henry P.
; REGISTRATION NUMBER: 33,200
; REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-127-721A-55

Query Match 73.2%; Score 41; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQNNWPPT 9
Db 1 QQSDSWPTT 9

RESULT 11
US-08-485-246A-55
; Sequence 55, Application US/08485246A
; Patent No. 6072035
; GENERAL INFORMATION:
; APPLICANT: Hardman, No. 6072035man
; APPLICANT: Kolbinger, Frank
; APPLICANT: Saldanha, Jose
; TITLE OF INVENTION: Reshaped monoclonal antibodies against an
; TITLE OF INVENTION: immunoglobulin isotype
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6072035artis Patent Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,246A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/127,721
; FILING DATE: 27-SEPTEMBER-1993
; APPLICATION NUMBER: US 07/952,802
; FILING DATE: 25-SEPTEMBER-1992
; NAME: No. 6072035ak, Henry P.
; REGISTRATION NUMBER: 33,200

```

REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-246A-55

Query Match 73.2%; Score 41; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNWPPT 9
DB 1 QQSDSWPT 9

RESULT 12

US-08-800-198-4
Sequence 4, Application US/08800198
Patent No. 5942602
GENERAL INFORMATION:
APPLICANT: WELS, WINFRIED S.
APPLICANT: SCHMIDT, MATHIAS
APPLICANT: VAKALOPOULOU, EVANGELIA
APPLICANT: SCHNEIDER, DOUGLAS
TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
STREET: 2200 CLARENDON BLVD. SUITE 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: US
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,198
FILING DATE: 13-FEB-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: HAMLET-KING, DIANA
REGISTRATION NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SCH 1576
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
US-08-800-198-4

Query Match 73.2%; Score 41; DB 2; Length 106;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNWPPT 9
DB 89 QQSDSWPT 97

RESULT 13
US-09-296-595-4
Sequence 4, Application US/09296595A
Patent No. 612915
GENERAL INFORMATION:
APPLICANT: WELS, WINFRIED S.
APPLICANT: SCHMIDT, MATHIAS
APPLICANT: VAKALOPOULOU, EVANGELIA
APPLICANT: SCHNEIDER, DOUGLAS
TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
FILE REFERENCE: SCH-1576 D1
CURRENT APPLICATION NUMBER: US/09/296,595A
CURRENT FILING DATE: 1999-04-23
EARLIER APPLICATION NUMBER: 08/800,198
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 106
TYPE: PRT
ORGANISM: Murine sp.
US-09-296-595-4

Query Match 73.2%; Score 41; DB 3; Length 106;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNWPPT 9
DB 89 QQSDSWPT 97

RESULT 14

US-08-476-176B-4
Sequence 4, Application US/08476176B
Patent No. 5958708
GENERAL INFORMATION:
APPLICANT: Hardman, No. 5958708man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958708artis Patent Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,176B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/127,721
FILING DATE: 27-SEPTEMBER-1993
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5958708ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-176B-4

Query Match 73.2%; Score 41; DB 2; Length 107;
Best Local Similarity 66.7%; Pred. NO. 12;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9
|:|:|:|:|
Db 89 QQSDSWPTT 97

RESULT 15
US-08-127-721A-4
Sequence 4, Application US/08127721A
Patent No. 6066718
GENERAL INFORMATION:
APPLICANT: Hardman, No. 6066718man
APPLICANT: Kolbinger, Frank
APPLICANT: Saldanha, Jose
TITLE OF INVENTION: Reshaped monoclonal antibodies against an
TITLE OF INVENTION: immunoglobulin isotype
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6066718artis Patent and Trademark Department
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: USA
ZIP: 07936-1080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,721A
FILING DATE: 27-SEPTEMBER-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,802
FILING DATE: 25-SEPTEMBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6066718ak, Henry P.
REGISTRATION NUMBER: 33,200
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-127-721A-4

Query Match 73.2%; Score 41; DB 3; Length 107;
Best Local Similarity 66.7%; Pred. NO. 12;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9
|:|:|:|:|
Db 89 QQSDSWPTT 97

Search completed: October 6, 2004, 16:36:34
Job time : 16.9474 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 16:20:48 ; Search time 45.4737 Seconds
(without alignments)
55.921 Million cell updates/sec

Title: US-09-635-974A-12
Perfect score: 56
Sequence: 1 QNNNNWPTT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|-------------|
| 1 | 56 | 100.0 | 9 | AAV59316 | Light cha |
| 2 | 56 | 100.0 | 9 | AAV59316 | Light cha |
| 3 | 56 | 100.0 | 9 | AAV59316 | Light cha |
| 4 | 56 | 100.0 | 107 | AAW08949 | Kappa lig |
| 5 | 56 | 100.0 | 107 | AAW08948 | Kappa lig |
| 6 | 56 | 100.0 | 127 | AAW08945 | Kappa lig |
| 7 | 56 | 100.0 | 127 | AAW08941 | Kappa lig |
| 8 | 56 | 100.0 | 127 | AAW08946 | Kappa lig |
| 9 | 56 | 100.0 | 127 | AAW08943 | Kappa lig |
| 10 | 56 | 100.0 | 240 | AAW05133 | Single ch |
| 11 | 56 | 100.0 | 651 | AAW05135 | scFv(225) |
| 12 | 56 | 100.0 | 892 | AAW05140 | scFv2(225) |
| 13 | 56 | 100.0 | 892 | AAW05139 | scFv2(FR3) |
| 14 | 56 | 100.0 | 1020 | AAW05141 | scFv2(FR3) |
| 15 | 45 | 80.4 | 107 | AAV32129 | Anti-IL2R |
| 16 | 45 | 80.4 | 240 | AAV34510 | Fv(TU27) |
| 17 | 44 | 78.6 | 224 | AAV99396 | Human int |
| 18 | 44 | 78.6 | 224 | AAV75029 | Anti-IL8 |
| 19 | 44 | 78.6 | 239 | AAV34511 | Fv(TU25) |
| 20 | 43 | 76.8 | 107 | AAV54897 | Light cha |
| 21 | 43 | 76.8 | 108 | AAV54906 | Light cha |
| 22 | 42 | 75.0 | 107 | AAV54897 | Humanised |
| 23 | 42 | 75.0 | 107 | AAV54895 | Light cha |
| 24 | 42 | 75.0 | 107 | AAV54896 | Light cha |
| 25 | 42 | 75.0 | 107 | AAV54891 | Light cha |

| | | | | | | |
|----|----|------|-----|---|----------|-----------|
| 26 | 42 | 75.0 | 107 | 6 | ABR54899 | Light cha |
| 27 | 42 | 75.0 | 107 | 6 | ABR54882 | Light cha |
| 28 | 42 | 75.0 | 107 | 6 | ABR54883 | Light cha |
| 29 | 42 | 75.0 | 107 | 6 | ABR54893 | Light cha |
| 30 | 42 | 75.0 | 107 | 7 | ADC61017 | Human ant |
| 31 | 42 | 75.0 | 107 | 7 | ADC61015 | Human ant |
| 32 | 42 | 75.0 | 107 | 7 | ADC99777 | Anti-huma |
| 33 | 42 | 75.0 | 107 | 7 | ADC99805 | Anti-huma |
| 34 | 42 | 75.0 | 107 | 7 | ADC99805 | Anti-huma |
| 35 | 42 | 75.0 | 107 | 7 | ADC99805 | Anti-MUC1 |
| 36 | 42 | 75.0 | 107 | 7 | ADC99805 | Anti-MUC1 |
| 37 | 42 | 75.0 | 112 | 7 | ADC61102 | Human ant |
| 38 | 42 | 75.0 | 127 | 7 | ADC61102 | Human ant |
| 39 | 42 | 75.0 | 127 | 7 | ADC61110 | Human ant |
| 40 | 42 | 75.0 | 134 | 4 | AAW77557 | Human bon |
| 41 | 42 | 75.0 | 134 | 4 | AAW77557 | Human liv |
| 42 | 42 | 75.0 | 134 | 5 | AAW77557 | Human pep |
| 43 | 42 | 75.0 | 246 | 5 | ABP45257 | Human Bly |
| 44 | 42 | 75.0 | 250 | 5 | ABP45941 | Human Bly |
| 45 | 42 | 75.0 | 498 | 7 | ADD13794 | Plasmid p |

ALIGNMENTS

RESULT 1
AAV59316
ID AAV59316 standard; peptide; 9 AA.

AC AAV59316;

XX 07-MAR-2000 (first entry)

XX Light chain hypervariable region, CDR3.

KW Hypervariable region; complementarity determining region; CDR; tumour;
KW single chain antibody; growth inhibitor; human; tumourigenesis; therapy;
KW protein receptor tyrosine kinase; light chain.

OS Mus sp.

XX WO9560023-A1.

XX 25-NOV-1999.

XX 14-MAY-1999; 99WO-US010741.

XX 15-MAY-1998; 98US-00079612.

XX 15-MAY-1998; 98US-0085613P.

XX 07-DEC-1998; 98US-00206138.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX (UABR-) UAB RES FOUND.

XX Wakeal HW, Saleh MN, Robert F, Buchsbaum DJ;

XX WPI, 2000-062440/05.

XX N-PSDB; AAZ48630.

XX Treatment of human tumors, using a combination of radiation and a non-

XX radiolabeled protein receptor tyrosine kinase inhibitor.

XX Disclosure; Page 15; 31pp; English.

CC This sequence is the hypervariable region CDR3 (complementarity
CC determining region 3) of the light chain of a single chain antibody
CC derived from the murine antibody 225. The invention relates to a method
CC for inhibiting the growth of tumours in human patients by treating with
CC an effective amount of a combination of radiation and a non-radiolabelled
CC protein receptor tyrosine kinase (PRTK) inhibitor, the overexpression of
CC which can lead to tumourigenesis. The method can be used in the treatment
CC of tumours of e.g. breast, lung, colon, kidney, bladder, head and neck,
CC ovary, prostate or brain. The administration of a suitable antibody to

CC the patient makes the tumour more susceptible to radiotherapy

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 56; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9

DB 1 QONNNWPTT 9

RESULT 2

AAB37957

ID AAB37957 standard; protein; 9 AA.

XX AC AAB37957;

XX DT 12-MAR-2001 (first entry)

XX DE Anti-EGFR monoclonal antibody I chain V region CDR3 peptide sequence.

XX KW Refractory tumour growth inhibition; epidermal growth factor receptor;
KW EGFR antagonist; cancer; squamous cell carcinoma; anti-EGFR antibody;
KW complementarity determining region; CDR.

XX OS Mus sp.

XX XN WO200069459-A1.

XX PD 23-NOV-2000.

XX PF 01-MAY-2000; 2000WO-US011756.

XX PR 14-MAY-1999; 99US-00312284.

XX PR 13-AUG-1999; 99US-00374028.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Wakeal HW;

XX DR WPI; 2001-016160/02.

XX DR N-PSDB; AAC83240.

XX PT Epidermal growth factor receptor/human epidermal growth factor receptor-1
antagonist for inhibiting the growth of refractory tumors.
XX PS Disclosure; Page 14; 31pp; English.

XX CC This invention relates to a method for inhibiting the growth of
CC refractory tumors that are stimulated by a ligand of epidermal growth
CC factor receptor (EGFR) in human patients. The method involves treating
CC the patient with a combination of EGFR/human EGF-1 (HER1) antagonist,
CC optionally with a chemotherapeutic agent or radiation. The antagonist can
CC be for example a chimeric anti-EGFR monoclonal antibody, C25. The
CC EGFR/HER1 antagonist is useful for inhibiting the growth of refractory
CC tumors such as tumors of breast, heart, lung, small intestine, colon,
CC spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas,
CC skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix and
CC liver, preferably squamous cell carcinomas. The present sequence
CC represents the light chain variable region complementarity determining
CC region 3 amino acid sequence of the chimeric anti-EGFR monoclonal
CC antibody C25 which is used in an example illustrating the method of the
CC invention

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 56; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9

DB 1 QONNNWPTT 9

RESULT 3

AAU77789

ID AAU77789 standard; peptide; 9 AA.

XX AC AAU77789;

XX DT 05-JUN-2002 (first entry)

XX DE Mouse light chain hypervariable region (CDR3) of 225 antibody.

XX KW Mouse; light chain; antibody; hyperproliferative disease;
KW epidermal growth factor; EGF; psoriasis; actinic keratosis;
KW seborrheic keratosis; warts; keloid scars; eczema; 255 antibody;
KW hypervariable region; CDR3; EGFR inhibitor.

XX OS Mus sp.

XX XN WO200211677-A2.

XX PD 14-FEB-2002.

XX PF 09-AUG-2001; 2001WO-US041647.

XX PR 09-AUG-2000; 2000US-00635974.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Teufel T;

XX DR WPI; 2002-257423/30.

XX DR N-PSDB; ABK11445.

XX PT Treating a mammal with hyperproliferative disease especially psoriasis,
PT stimulated by ligand of member of epidermal growth factor family of
PT receptors, by administering antagonist of the receptor.
XX PS Disclosure; Page 12; 28pp; English.

XX CC This invention relates to a novel method for treating a mammal with
CC hyperproliferative disease stimulated by a ligand of a member of the
CC epidermal growth factor (EGF) family of receptors. The method involves
CC administering an antibody or a defective receptor that is an antagonist
CC of a member of the EGF receptor family, or a combination of the
CC antagonist and phototherapy, chemotherapeutic agent or radiation therapy.
CC The antibody used in the method of the invention acts as an epidermal
CC growth factor receptor (EGFR) antagonist by inhibiting EGFR/HER1
CC phosphorylation. The method of the invention is useful for treating a
CC mammal with hyperproliferative disease such as psoriasis, actinic
CC keratoses, seborrheic keratoses, warts, keloid scars and eczema
CC stimulated by a ligand of a member of the EGF family of receptor. This
CC sequence represents the murine anti-EGF 255 antibody light chain (LH)
CC hypervariable region (CDR3) used as an inhibitor of EGFR in the method of
CC the invention

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 56; DB 5; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QONNNWPTT 9

DB 1 QONNNWPTT 9

RESULT 4

AAW08949

ID AAW08949 standard; protein; 107 AA.

XX

```

AC AAW08949;
XX
DT 18-SEP-1997 (first entry)
XX
DE Kappa light chain variable region of 225RB antibody.
XX
KW Kappa; light chain; reshaped; monoclonal; antibody; 225RB; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 1..23
FT /label= framework_1
FT Region 24..34
FT /label= CDR_1
FT Region 35..49
FT /label= framework_2
FT Region 50..56
FT /label= CDR_2
FT Region 57..88
FT /label= framework_3
FT Region 89..97
FT /label= CDR_3
FT Region 98..107
FT /label= framework_4
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
DR WPI; 1997-051897/05.
XX
PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PS Claim 29; Fig 21; 112pp; English.
XX
CC The present sequence is the kappa light chain variable region of the
CC reshaped human monoclonal antibody (Mab) H225, 225RKA. The Mab is
CC specific for the human epidermal growth factor (EGF) receptor. The Mab,
CC or a fragment, can be used to inhibit the growth of tumour cells,
CC especially late stage prostatic tumour cells in humans, optionally
CC conjugated to a cytotoxic agent, especially doxorubicin, taxol or
CC cisplatin, or a signal transduction, ras or cell cycle inhibitor
XX
SQ Sequence 107 AA;
Query Match 100.0%; Score 56; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQNNNWPTT 9
Db 89 QQNNNWPTT 97
RESULT 5
AAW08948
ID AAW08948 standard; protein; 107 AA.
XX

```

```

AC AAW08948;
XX
DT 18-SEP-1997 (first entry)
XX
DE Kappa light chain variable region of 225RA antibody.
XX
KW Kappa; light chain; reshaped; monoclonal; antibody; 225RA; human;
KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
KW late stage; prostatic; prostate; variable region; framework;
KW complementarity determining region; CDR.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Region 1..23
FT /label= framework_1
FT Region 24..34
FT /label= CDR_1
FT Region 35..49
FT /label= framework_2
FT Region 50..56
FT /label= CDR_2
FT Region 57..88
FT /label= framework_3
FT Region 89..97
FT /label= CDR_3
FT Region 98..107
FT /label= framework_4
XX
PN WO9640210-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009847.
XX
PR 07-JUN-1995; 95US-00482982.
PR 15-DEC-1995; 95US-00573289.
XX
PA (IMCL-) IMCLONE SYSTEMS INC.
PA (MRCC-) MRC COLLABORATIVE CENT.
XX
PI Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
DR WPI; 1997-051897/05.
XX
PT Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
PS Claim 29; Fig 21; 112pp; English.
XX
CC The present sequence is the kappa light chain variable region of the
CC reshaped human monoclonal antibody (Mab) H225, 225RKA. The Mab is
CC specific for the human epidermal growth factor (EGF) receptor. The Mab,
CC or a fragment, can be used to inhibit the growth of tumour cells,
CC especially late stage prostatic tumour cells in humans, optionally
CC conjugated to a cytotoxic agent, especially doxorubicin, taxol or
CC cisplatin, or a signal transduction, ras or cell cycle inhibitor
XX
SQ Sequence 107 AA;
Query Match 100.0%; Score 56; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QQNNNWPTT 9
Db 89 QQNNNWPTT 97
RESULT 6
AAW08945
ID AAW08945 standard; protein; 127 AA.
XX

```

AC AAW08945;
 XX 18-SEP-1997 (first entry)
 DT Kappa light chain variable region of C225 antibody.
 DE Kappa; light chain; murine; mouse; monoclonal; antibody; C225; human;
 XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 KW late stage; prostate; prostate; variable region; chimeric; L7'CL;
 KW leader sequence.
 XX Mus; spp.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 XX WO9640210-A1.
 PN 19-DEC-1996.
 PD 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 PR 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 DR N-PSDB; AAT49344.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Example IV; Fig 17; 112pp; English.
 XX The present sequence is the kappa light chain variable region of the
 CC chimeric monoclonal antibody (Mab) C225, with the modified leader
 CC sequence from the kappa light chain of L7'CL Mab. C225 is specific for
 CC the human epidermal growth factor (EGF) receptor. C225, or a fragment,
 CC can be used to inhibit the growth of tumour cells, especially late stage
 CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 CC agent, especially doxorubicin, taxol or cisplatin, or a signal
 CC transduction, ras or cell cycle inhibitor
 XX Sequence 127 AA;
 SQ Query Match 100.0%; Score 56; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.099;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QQNNNWPTT 9
 DB 109 QQNNNWPTT 117
 RESULT 7
 AAW08941
 ID AAW08941 standard; protein; 127 AA.
 AC AAW08941;
 XX 27-AUG-2003 (revised)
 DT 18-SEP-1997 (first entry)
 DE Kappa light chain variable region of M225 antibody.
 XX Kappa; light chain; murine; mouse; monoclonal; antibody; M225; human;
 KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 KW late stage; prostate; prostate; variable region.
 XX Mus; spp.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 XX WO9640210-A1.
 PN 19-DEC-1996.
 PD 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 PR 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 DR N-PSDB; AAT49344.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Example IV; Fig 17; 112pp; English.
 XX The present sequence is the kappa light chain variable region of the
 CC chimeric monoclonal antibody (Mab) C225, with the modified leader
 CC sequence from the kappa light chain of L7'CL Mab. C225 is specific for
 CC the human epidermal growth factor (EGF) receptor. C225, or a fragment,
 CC can be used to inhibit the growth of tumour cells, especially late stage
 CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 CC agent, especially doxorubicin, taxol or cisplatin, or a signal
 CC transduction, ras or cell cycle inhibitor
 XX Sequence 127 AA;
 SQ Query Match 100.0%; Score 56; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.099;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QQNNNWPTT 9
 DB 109 QQNNNWPTT 117
 RESULT 7
 AAW08941
 ID AAW08941 standard; protein; 127 AA.
 AC AAW08941;
 XX 27-AUG-2003 (revised)
 DT 18-SEP-1997 (first entry)
 DE Kappa light chain variable region of M225 antibody.
 XX Kappa; light chain; murine; mouse; monoclonal; antibody; M225; human;
 KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 KW late stage; prostate; prostate; variable region.

OS Mus sp.
 XX WO9640210-A1.
 PN 19-DEC-1996.
 PD 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 PR 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 XX WPI; 1997-051897/05.
 DR N-PSDB; AAT49338.
 XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
 PT for inhibiting tumour growth, esp. of late stage prostatic tumour.
 XX Example IV; Fig 13; 112pp; English.
 XX The present sequence is the kappa light chain variable region of the
 CC murine monoclonal antibody (Mab) M225, which is specific for the human
 CC epidermal growth factor (EGF) receptor. The Mab, or a fragment, can be
 CC used to inhibit the growth of tumour cells, especially late stage
 CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
 CC agent, especially doxorubicin, taxol or cisplatin, or a signal
 CC transduction, ras or cell cycle inhibitor. (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX Sequence 127 AA;
 SQ Query Match 100.0%; Score 56; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 0.099;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QQNNNWPTT 9
 DB 109 QQNNNWPTT 117
 RESULT 8
 AAW08946
 ID AAW08946 standard; protein; 127 AA.
 AC AAW08946;
 XX 18-SEP-1997 (first entry)
 DT Kappa light chain variable region of 225RA antibody.
 DE Kappa; light chain; reshaped; monoclonal; antibody; 225RA; human;
 KW epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
 KW late stage; prostate; prostate; variable region.
 XX Homo sapiens.
 OS WO9640210-A1.
 PN 19-DEC-1996.
 PD 07-JUN-1996; 96WO-US009847.
 XX 07-JUN-1995; 95US-00482982.
 PR 15-DEC-1995; 95US-00573289.
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA (MRCC-) MRC COLLABORATIVE CENT.
 XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
 PI

```

XX WPI; 1997-051897/05.
DR N-PSDB; AAT49345.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Claim 31; Fig 19; 112pp; English.
XX
XX The present sequence is the kappa light chain variable region of the
XX reshaped human monoclonal antibody (MAB) H225, 225RKA. The MAB is
XX specific for the human epidermal growth factor (EGF) receptor. The MAB,
XX or a fragment, can be used to inhibit the growth of tumour cells,
XX especially late stage prostatic tumour cells in humans, optionally
XX conjugated to a cytotoxic agent, especially doxorubicin, taxol or
XX cisplatin, or a signal transduction, ras or cell cycle inhibitor
XX
XX Sequence 127 AA;
XX
XX Query Match 100.0%; Score 56; DB 2; Length 127;
XX Best Local Similarity 100.0%; Pred. No. 0.099;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 QONNNWPTT 9
XX |||||
XX Db 109 QONNNWPTT 117
XX
XX RESULT 9
XX AAW08943
XX ID AAW08943 standard; protein; 127 AA.
XX AC AAW08943;
XX
XX DT 18-SEP-1997 (first entry)
XX
XX DE Kappa light chain variable region of C225 antibody.
XX
XX Kappa; light chain; murine; mouse; monoclonal; antibody; C225; human;
XX epidermal growth factor; EGF; receptor; inhibition; growth; tumour; cell;
XX late stage; prostatic; prostate; variable region; chimeric.
XX
XX Mus; spp.
XX Homo sapiens.
XX Synthetic.
XX Chimeric.
XX
XX WO9640210-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009847.
XX
XX 07-JUN-1995; 95US-00482982.
XX 15-DEC-1995; 95US-00573289.
XX
XX (IMCL-) IMCLONE SYSTEMS INC.
XX (MRCC-) MRC COLLABORATIVE CENT.
XX
XX Goldstein NI, Giorgio NA, Jones ST, Saldanha JW;
XX
XX WPI; 1997-051897/05.
XX N-PSDB; AAT49342.
XX
XX Chimeric and humanised versions of anti-EGF receptor antibody 225 - used
XX for inhibiting tumour growth, esp. of late stage prostatic tumour.
XX
XX Example IV; Fig 15; 112pp; English.
XX
XX The present sequence is the kappa light chain variable region of the
XX chimeric monoclonal antibody (MAB) C225, which is specific for the human
XX epidermal growth factor (EGF) receptor. The MAB, or a fragment, can be
XX used to inhibit the growth of tumour cells, especially late stage

```

```

CC prostatic tumour cells in humans, optionally conjugated to a cytotoxic
CC agent, especially doxorubicin, taxol or cisplatin, or a signal
CC transduction, ras or cell cycle inhibitor
XX
XX Sequence 127 AA;
XX
XX Query Match 100.0%; Score 56; DB 2; Length 127;
XX Best Local Similarity 100.0%; Pred. No. 0.099;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 QONNNWPTT 9
XX |||||
XX Db 109 QONNNWPTT 117
XX
XX RESULT 10
XX AAW05133
XX ID AAW05133 standard; protein; 240 AA.
XX AC AAW05133;
XX
XX DT 29-JAN-1997 (first entry)
XX
XX DE Single chain antibody scFv(225).
XX
XX Single chain antibody; scFv; monoclonal antibody; MAB; EGF;
XX epidermal growth factor; receptor; antitumour; cancer; therapy.
XX
XX Mus; sp.
XX Synthetic.
XX Chimeric.
XX
XX FH Key Location/Qualifiers
XX FT Region 1..119
XX FT /label= VH region
XX FT /note= "monoclonal antibody 225 VH"
XX FT Peptide 120..133
XX FT /label= Linker
XX FT /note= "synthetic spacer peptide"
XX FT Region 134..240
XX FT /label= VL
XX FT /note= "monoclonal antibody 225 VL"
XX
XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-478748/48.
XX N-PSDB; AAT42033.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 2; Page 17; 52pp; English.
XX
XX scFv(225) (AAW05133) comprises the single-chain binding region of murine
XX monoclonal antibody 225, which is specific for the human epidermal growth
XX factor receptor. It is encoded by plasmid pMW152-225 (see also AAT42033),
XX constructed by cloning MAB 225 VH and VL region cDNAs into plasmid
XX pMW152. Novel bivalent proteins (see also AAW05134-44), some of them
XX including scFv(225) and an effector e.g. cytotoxin, can be produced in
XX bacterial host cells, and are useful as antitumour agents
XX
XX Sequence 240 AA;
XX

```

Query Match 100.0%; Score 56; DB 2; Length 240;

Best Local Similarity 100.0%; Pred. No. 0.19; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9

Db 223 QNNNNWPTT 231

RESULT 11

AAW05135
ID AAW05135 standard; protein; 651 AA.

XX AC AAW05135;

XX DT 29-JAN-1997 (first entry)

XX DE scFv(225)-ETA fusion protein.

XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF;

XX KW epidermal growth factor; receptor; plasmid pSW202-225; cancer; therapy;

XX KW antitumour; exotoxin A; ETA.

XX OS Mus sp.

XX OS Pseudomonas; aeruginosa.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key

XX FT Peptide

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

SQ Sequence 651 AA;

Query Match 100.0%; Score 56; DB 2; Length 651;

Best Local Similarity 100.0%; Pred. No. 0.55; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9

Db 261 QNNNNWPTT 269

RESULT 12

AAW05140

ID AAW05140 standard; protein; 892 AA.

XX AC AAW05140;

XX DT 29-JAN-1997 (first entry)

XX DE scFv(225)/FRP5)-ETA.

XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;

XX KW epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;

XX KW exotoxin A; ETA; antitumour.

XX OS Mus, sp.

XX OS Pseudomonas; aeruginosa.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key

XX FT Peptide

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

Query Match 100.0%; Score 56; DB 2; Length 240;

Best Local Similarity 100.0%; Pred. No. 0.19; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9

Db 223 QNNNNWPTT 231

RESULT 11

AAW05135
ID AAW05135 standard; protein; 651 AA.

XX AC AAW05135;

XX DT 29-JAN-1997 (first entry)

XX DE scFv(225)-ETA fusion protein.

XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF;

XX KW epidermal growth factor; receptor; plasmid pSW202-225; cancer; therapy;

XX KW antitumour; exotoxin A; ETA.

XX OS Mus sp.

XX OS Pseudomonas; aeruginosa.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key

XX FT Peptide

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

SQ Sequence 651 AA;

Query Match 100.0%; Score 56; DB 2; Length 651;

Best Local Similarity 100.0%; Pred. No. 0.55; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9

Db 261 QNNNNWPTT 269

RESULT 12

AAW05140

ID AAW05140 standard; protein; 892 AA.

XX AC AAW05140;

XX DT 29-JAN-1997 (first entry)

XX DE scFv(225)/FRP5)-ETA.

XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;

XX KW epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;

XX KW exotoxin A; ETA; antitumour.

XX OS Mus, sp.

XX OS Pseudomonas; aeruginosa.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key

XX FT Peptide

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Peptide

XX FT Peptide

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

XX FT Protein

SQ Sequence 651 AA;

Query Match 100.0%; Score 56; DB 2; Length 651;

Best Local Similarity 100.0%; Pred. No. 0.55; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QNNNNWPTT 9

Db 261 QNNNNWPTT 269

RESULT 12

AAW05140

ID AAW05140 standard; protein; 892 AA.

XX AC AAW05140;

XX DT 29-JAN-1997 (first entry)

XX DE scFv(225)/FRP5)-ETA.

XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;

XX KW epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;

XX KW exotoxin A; ETA;

XX scFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region
 CC of murine monoclonal antibody 225 (specific for human epidermal growth
 CC factor receptor, see also AAW05133) joined to portions of exotoxin A from
 CC Pseudomonas aeruginosa and to the single-chain binding region of murine
 CC monoclonal antibody FRP5 (specific for human epidermal growth factor
 CC receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-
 CC 5 (AA42040). This plasmid can be utilised in the prodn. of the bivalent
 CC fusion protein in bacterial (esp. E. coli) host cells. Such fusion
 CC proteins (see also AAW05138-44) are useful as antitumour agents
 XX
 SQ Sequence 892 AA;

Query Match 100.0%; Score 56; DB 2; Length 892;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQNNNWPTT 9
 |||||
 Db 261 QQNNNWPTT 269

RESULT 13
 AAW05139
 ID AAW05139 standard; protein; 892 AA.
 XX AC AAW05139;

XX 29-JAN-1997 (first entry)
 XX DE scFv2(FRP5/225)-ETA (version 1).
 XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;
 KW epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
 KW exotoxin A; ETA; antitumour.
 XX OS Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT /note= "ompA signal peptide"
 FT Peptide 22..38
 FT /label= Spacer
 FT Region 39..278
 FT /label= scFv(FRP5)
 FT Peptide 279..289
 FT /label= Spacer
 FT Region 290..404
 FT /label= ETA
 FT Peptide 405..407
 FT /note= "exotoxin A amino acids 252-366"
 FT Region 408..647
 FT /label= scFv(225)
 FT Peptide 648..658
 FT /label= Spacer
 FT Region 659..892
 FT /label= ETA
 FT /note= "endotoxin-A amino acids 380-613"

XX EP739984-A1.
 XX 30-OCT-1996.
 XX 26-APR-1995; 95EP-00106275.
 XX 26-APR-1995; 95EP-00106275.
 XX (SANT-) SAN TUMORFORSCHUNGS GMBH.

XX Wells W. Schmidt M, Groner B;
 XX WPI: 1996-478748/48.
 DR N-PEDB; AAR42039.
 XX Bivalent fusion proteins that bind epidermal growth factor receptor or
 PT analogues - and comprise at least two different cell surface binding
 PT domain(s), useful for tumour therapy.
 XX Example 11; Page 28-30; 52pp; English.

XX scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
 CC of murine monoclonal antibody FRP5 (specific for human epidermal growth
 CC factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
 CC A from Pseudomonas aeruginosa and to the single-chain binding region of
 CC murine monoclonal antibody 225 (specific for human epidermal growth
 CC factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
 CC 225 (AA42039). This plasmid can be utilised in the prodn. of the
 CC bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
 CC fusion proteins (see also AAW05138-44) are useful as antitumour agents
 XX
 SQ Sequence 892 AA;

Query Match 100.0%; Score 56; DB 2; Length 892;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQNNNWPTT 9
 |||||
 Db 630 QQNNNWPTT 638

RESULT 14
 AAW05141
 ID AAW05141 standard; protein; 1020 AA.

XX AC AAW05141;
 XX DT 29-JAN-1997 (first entry)
 XX DE scFv2(FRP5/225)-ETA (version 2).
 XX KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;
 KW epidermal growth factor; receptor; plasmid pMS240-5-225; cancer;
 KW exotoxin A; ETA; antitumour.
 XX OS Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT /note= "ompA signal peptide"
 FT Peptide 22..38
 FT /label= Spacer
 FT Region 39..278
 FT /label= scFv(FRP5)
 FT Peptide 279..289
 FT /label= Spacer
 FT Region 290..404
 FT /label= ETA
 FT /note= "exotoxin A amino acids 252-366"
 FT Peptide 405..407
 FT /label= Spacer
 FT Region 408..647
 FT /label= scFv(225)
 FT Peptide 648..658
 FT /label= Spacer
 FT Region 659..1020
 FT /label= ETA

